The Hague Centre
for Strategic Studies

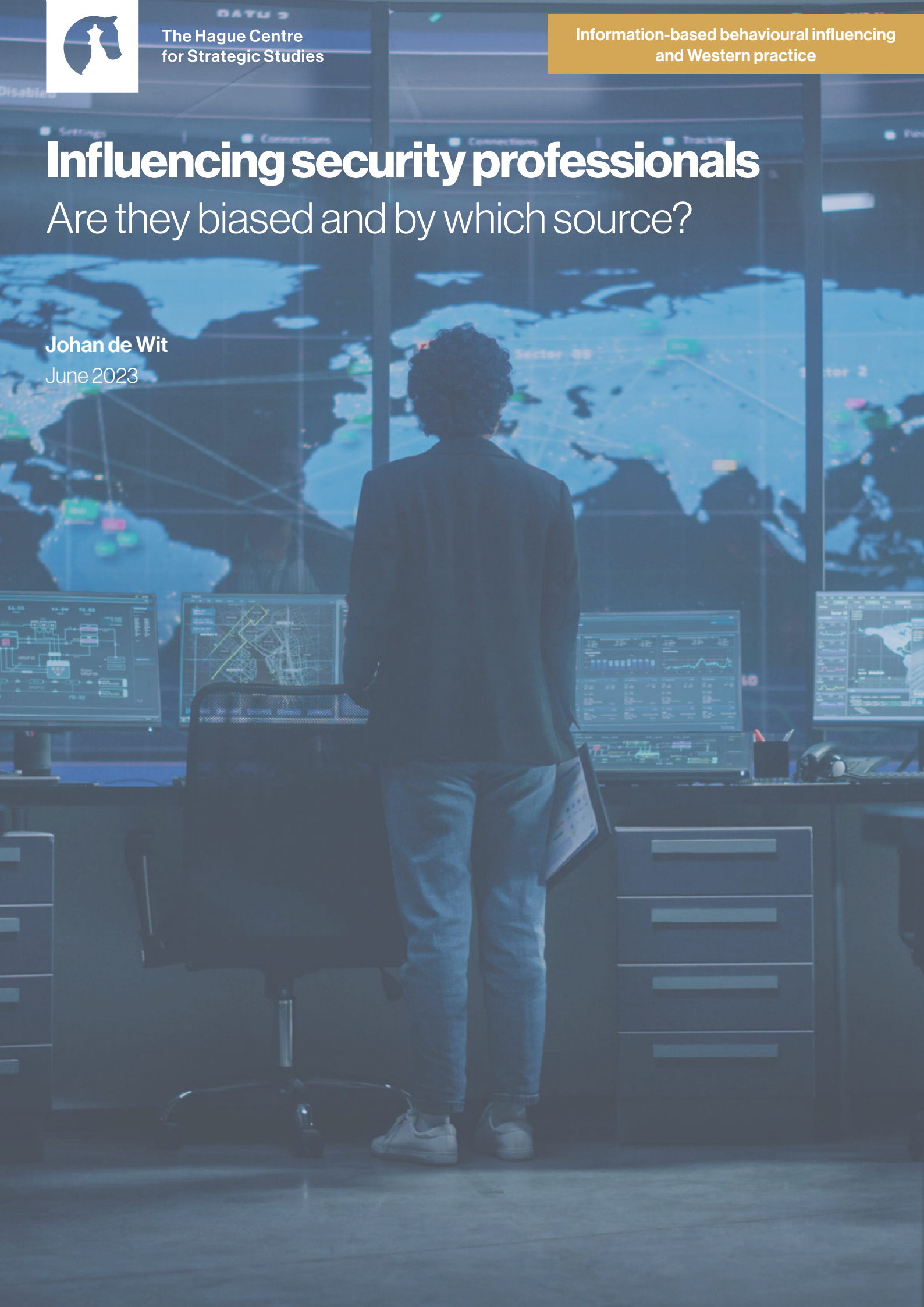
Information-based behavioural influencing
and Western practice

Influencing security professionals

Are they biased and by which source?

Johan de Wit

June 2023



The Hague Centre
for Strategic Studies

Paper 5

Influencing security professionals

Are they biased and by which source?

Author:
Johan de Wit

This paper is part of the *Information-based behavioural influencing and Western practice* paper series.

June 2023

This paper is published as part of the project Platform Influencing Human Behaviour, commissioned by the Royal Netherlands Army. The aim of this platform is to build and share knowledge on information-based behavioural influencing in the military context. We bring together international experts and practitioners from both military and academic backgrounds to explore the military-strategic, ethical, legal, and societal issues and boundaries involved. Responsibility for the content rests solely with the authors and does not constitute, nor should it be construed as, an endorsement by the Royal Netherlands Army.

Paper series: Information-based behavioural influencing and Western practice

The military application of information has a long history in influencing the outcome of war and conflict on the battlefield. Be it by deceiving the opponent, maintaining troop confidence, or shaping public opinion. These tactics are placed under the banner of influencing human behaviour. Behavioural influencing is the act of meaningfully trying to affect the behaviour of an individual by targeting people's knowledge, beliefs and emotions. Within the Dutch armed forces these tactics fall under title of Information Manoeuvre. With the ever-larger and more evasive employment of information-based capabilities to target human cognition, the boundaries of the physical and cognitive battlefield have begun to fade.

This paper is published as part of the project *Platform Influencing Human Behaviour*, commissioned by the Royal Netherlands Army. The aim of this platform is to build and share knowledge on information-based behavioural influencing in the military context. We bring together international experts and practitioners from both military and academic backgrounds to explore the military-strategic, ethical, legal, and societal issues and boundaries involved. Responsibility for the content rests solely with the authors and does not constitute, nor should it be construed as, an endorsement by the Royal Netherlands Army.

For this paper series scholars, experts and policymakers submitted their papers on the employment of information-related capabilities to influence human behaviour in the military context. From the perspective of an individual European or NATO country's perspective. The Information-based behavioural influencing and Western practice paper series is edited by Arthur Laudrain, Laura Jasper and Michel Rademaker.

Seven papers will be published in this series. These are the following:

- **Deception as the Way of Warfare. Armed Forces, Influence Operations and the Cyberspace paradox.** By Colonel dr. Peter B.M.J. Pijpers, Netherlands Defence Academy and University if Amsterdam, and Brigadier-General prof. dr. Paul A.L. Ducheine, Netherlands Defence Academy and University of Amsterdam
- **Influencing security professionals: are they biased and by which source?** By Johan de Wit, TU Delft & Siemens Smart Infrastructure
- **A discursive analytical approach to understanding target audiences. How NATO can improve its actor-centric analysis.** By Yannick Smits, Research Master Middle Eastern studies Leiden University
- **The concept of Information Manoeuvre: Winning the Battle of Perceptions.** By Judith T. van de Kuijt (TNO), N. Keja (TNO), J.C. Slaager (TNO)
- **Smart Tactics or Risky Behaviour? The Lawfulness of Encouraging Civilians to Participate in Targeting in an Age of Digital Warfare.** By Pontus Winther, LL.D. Swedish Armed Forces, and Per-Erik Nilsson, Ph.D. Swedish Defence Research Agency and Associate Professor at Uppsala University
- **Cognitive Warfare as Part of Society: Never-Ending Battle for Minds.** By Robin Burda, Ph.D. candidate Security and Strategic Studies Masaryk University
- **Behavioural Influence Interventions in the Information Environment: Underlying Mechanisms and Technologies.** By dr. Hans Korteling (TNO), Beatrice Cadet (TNO), Tineke Hof (TNO)

Abstract

This paper presents brief summaries of four studies that explore the factors that drive our intuitive or reasoned perceptions of risk. The first part presents two studies on information and the sources of information that are the foundations for this risk perception. The second part of this paper presents the summaries of two studies that explore the biases and heuristics that affect the decision maker in the interpretation of information. These studies are all conducted in the professional security domain to investigate real-life security risk decision making. The summaries in this paper do not include extensive methods and analysis sections, we kindly refer to the published full papers. The results this paper identify some fundamental human traits that can be exploited to influence human decision behaviour. On the other hand, any responsible decision maker should be aware of them and take them into account in their own daily praxis, as the results clearly and undoubtedly show the effects of these phenomena on judgements of, especially, experienced professionals.

While risk is often portrayed mathematically, our response is more often instinctive. Understanding the factors that drive how we think about and act upon risk is critical.

General Stanley McChrystal, US Army, retired[1]

Introduction

In the domain of security risks, which is dealing with risks originating from malicious human action, risk assessments are predominantly based on expert judgment. [2] Although information on threats and risks might be available, it is often incomplete and imperfect so expert interpretation is usually influencing and/or decisive for a security risk assessment.

Human decision making is prone to heuristics and biases as decades of scientific work demonstrated.[3 4 5] Our work, as part of a PhD research project, studies the influence of heuristics and biases on individual risk decision making and risk assessments. The influence of information, triggering heuristics and biases, is the cornerstone of our studies. Several of them, both published and to be published, presented evidence of the influence of biases and heuristics on risk decision making and assessments by security practitioners. Our research has shown that:

- biases lead to less effective security decisions by professionals,
- risk attribute preferences can be influenced,
- more detailed information raises the likelihood perception,
- the widespread existence of probability ignorance in security risk decision making,
- security decision makers show objective ignorance,
- security decision makers are notorious overconfident even if they are aware their information is incomplete and imperfect

Our work may not directly answer the research questions and issues as posed for the Platform Influencing Human Behaviour, and is not specifically addressing decision making in

1 Stanley A. McChrystal, *Risk, a user's guide* (New York: Penguin business, 2021).

2 Niklas Möller, "The concepts of risk and safety," in *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk* (Springer, 2012).

3 Daniel Kahneman and Amos Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica: Journal of the econometric society* (1979).

4 Herbert Alexander. Simon, *Models of bounded rationality: Empirically grounded economic reason*, vol. 3 (MIT press, 1982).

5 Gerd Gigerenzer, *Risk savvy: How to make good decisions* (Penguin, 2015).

a military context. Our studies might, however, contribute some valuable insights in individual risk decision making and risk assessments. This perspective can help forming the needed policies for influencing individual security assessments of military decision makers.

This paper presents brief summaries of four different studies of security risk assessments in the security domain by individual risk professionals. They address two main topics:

1. the information on which assessments are based (identify sources, how much security risk information is available, how does this influence confidence)
2. biases and heuristics influencing the interpretation and perception of this information (study of vulnerability for known biases, conjunction fallacy, availability/on top-of-mind study, system 1 and 2 thinking)

In the next section the results of two studies are presented that explore security risk information. Both studies are briefly introduced and the relevant results are presented. The following section contains two studies into the vulnerability for, and influence of, biases on security risk decision making. This paper ends with some overall conclusions. In the summaries of the studies a research method section is deliberately left out to make this paper fit the maximum size. We kindly refer to the full papers for an extensive overview of the methods and analysis.

Topic 1: information, the foundation of risk assessments

In this section we present a summary of the results of two studies. In the first study we explore the information position of security professionals (the level of availability of precise information and/or evidence). We investigate how this position influences the confidence in their own judgment. The influence of individual expertise, on both the need for information and confidence levels, is examined. This study is published in a full paper: "Bias and Noise in Security Risk Assessments, an Empirical Study on the Information Position and Confidence of Security Professionals", to be published in: *Security Journal*. The second study collects the possible sources of security risk information. It explores both the perceived quality and trustworthiness, and the application in real life of information sources. This study is published in a full paper: "Sources of Security Risk Information: What do Professionals Rely on for their Risk Assessment?", currently under review.

The first empirical study addresses the following research questions:

- Do security professionals have exact information on security risks during their risk assessments?
- How confident are they about their security risk assessment?
- Would more information grow their confidence?

As the future cannot be certain by nature, professionals might be expected to 'know that they cannot possibly know' (known unknowns).

The results of a survey on the information position of security professionals are presented in Table 1. The professionals are asked to indicate, based on their real life praxis, the level of detail of security risk information available to them. The study focusses on the two main components of risk: impact and uncertainty expressed as likelihood. The security professionals indicate that, on average, about half the time they know the likelihood and consequences of the security risks they are assessing exactly. The respondents also indicate that they, on average, only sometimes, cannot estimate likelihood and consequences. One in four even indicates that they can always estimate likelihood and consequences, based on their experience and knowledge, even when they indicate they know they do not have accurate information.

This finding deviates from the expectation that security professionals would recognize their information position about security risks as both imperfect and intractable. As the future cannot be certain by nature, professionals might be expected to 'know that they cannot possibly know' (known unknowns).[6] They, however, indicate that they can estimate the likelihood and consequences most of the time. Assuming that the respondents are right about their knowledge position, they assess risks half of the time based on information (evidence based). On the other hand they assess security risks without proper information also half of the time and still come up with an estimation of likelihood and consequences. As these assessments have a serious impact on security risk decision making and the allocation of resources to manage, mitigate and/or accept these risks, it is worth noting that these decisions don't seem to be based on evidence about half of the time.

The perception of the respondents on their information position can be questioned. As risk assessments are in fact predictive judgements and the information about the future can be considered intractable by nature, this perception of the security professionals can be considered audacious.

6 Daniel Kahneman, Olivier Sibony, and Cass R. Sunstein, *Noise, a Flaw in Human Judgment* (London: William Collins, 2021).

Table 1. The information position of security professionals in security risk assessments.

When evaluating security risks in general:	Always (1)	Most of the time (2)	About half the time (3)	Sometimes (4)	Never (5)	Median answer	Mean answer*
I know the **likelihood** of security events **exactly**	2.0%	33.0%	19.3%	24.9%	20.8%	About half the time	3.29
I do not know the **likelihood** exactly but I have **quantified** information (evidence based probability)	4.6%	38.1%	23.4%	29.9%	4.1%	About half the time	2.91
I do not know the **likelihood** exactly but I **can estimate** the likelihood based on my experience and knowledge	9.6%	51.3%	23.9%	14.7%	0.5%	Most of the time	2.45
I do not know the **likelihood** exactly and I **cannot estimate** the likelihood based on my experience and knowledge	0.5%	13.7%	8.1%	54.3%	23.4%	Sometimes	3.86
I know the **consequences** of security events **exactly**	3.3%	42.4%	21.2%	19.6%	13.6%	About half the time	2.98
I do not know the **consequences** exactly but I have **quantified** information (evidence based probability)	3.3%	39.7%	20.7%	31.5%	4.9%	About half the time	2.95
I do not know the **consequences** exactly but I **can estimate** the likelihood based on my experience and knowledge	7.1%	49.5%	21.7%	19.6%	2.2%	Most of the time	2.60
I do not know the **consequences** exactly and I **cannot estimate** the likelihood based on my experience and knowledge	0.5%	12.0%	8.7%	50.5%	28.3%	Sometimes	3.94

** considering the Likert scale a continues variable from always = 1 to never = 5*

Overall, the respondents claim to be confident about their judgement of likelihood and consequences most of the time (see Table 2).

Table 2. Confidence levels of security professionals.

When evaluating security risks in general:	Always (1)	Most of the time (2)	About half the time (3)	Sometimes (4)	Never (5)	Median answer	Mean answer*
I feel confident about my assessments of the **likelihood** of security risks	8.3%	59.4%	20.0%	10.6%	1.7%	Most of the time	2.38
I feel confident about my assessments of the **consequences** of security risks	9.4%	64.4%	15.6%	9.4%	1.1%	Most of the time	2.28
I would feel **more confident** if I had more information on security risks	28.9%	33.3%	8.9%	27.8%	1.1%	Most of the time	2.39

** considering the Likert scale a continues variable from always = 1 to never = 5*

Again: as the future cannot be certain by nature, the confidence of security professionals in their predictive judgements is expected to be limited. Overall the majority of the security professionals, however, indicate that they are always or most of the time confident about their assessments. It is hypothesized that the security professionals would show modest confidence based on the assumption that exact and/or evidence based information on security risks is often lacking. They, however, seem to ignore the latter and thus show a higher level of confidence than might be expected. As the respondents, on average, indicate to hold exact or quantified information only half of the time, they, thus, might be considered overconfident about their risk assessments.

Combining the perceived information position of the professionals with their level of confidence reveals objective ignorance.[7] A portion of respondents indicate they have exact information only sometimes or even never but are confident most or half of the time. These respondents are aware of their lack of exact information but are confident nevertheless. This lack of information doesn't seem to affect their ability to form a predictive judgement and be confident about it.

In this study the respondents are asked to indicate their age, number of years professional and security experience, their general education level (associate degree, bachelor degree or master degree/PhD) and if any specific security trainings are completed.

This lack of information doesn't seem to affect their ability to form a predictive judgement and be confident about it.

More professional and security experience significantly raises the confidence level of the security professionals. More experienced security professionals are more often confident about their assessments of both likelihood and consequences. More experienced security professionals also indicate that more information would raise their confidence level to a lesser extent than less experienced professionals indicate. In short these results seem to indicate that more experience leads to higher levels of (over)confidence and less need for additional information. A higher education level on the other hand significantly reduces the confidence in likelihood and consequences assessments. These results might prove the adage 'the more you know, the more you realise you don't know' as other scholars also found. [8] Security specific trainings do not significantly influence confidence level or the need for additional information.

In the second study the origin of security risk information, the sources of information, are explored and studied. Possible sources of security risk information are collected, their quality and trustworthiness are assessed, and the level of their application in real life security risk assessments is analysed. The research questions answered in this study are:

- What sources of security risk information are considered by practitioners?
- How reliable are these sources as perceived by these practitioners?
- Which sources are applied in security risk assessment praxis?
- Are the most applied sources also perceived as the most credible ones?
- Can we observe differences between security professionals based on their expertise (experience and knowledge)?

Analysing and classifying information and information sources is of vital importance in the security domain. [9] [10] [11] Especially in the security intelligence community tools and methods are developed and applied to classify information and information sources. [12] [13] In this domain the quality of information is also predominantly evaluated based on both the reliability of the content and the source, applying the international and broadly accepted evaluation criteria known as the Admiralty Code or NATO System (see Table 3). The NATO system classifies the reliability of sources on: authenticity, trustworthiness and competency. These characteristics are evaluated against past experience with the sources.

7 Cass R Sunstein, *Laws of fear: beyond the precautionary principle* (Cambridge University Press, 2005).

8 George Wright and Peter Ayton, "Subjective confidence in forecasts: A response to Fischhoff and MacGregor," *Journal of Forecasting* 5, no. 2 (1986).

9 Thomas Powell et al., "Dealing with Uncertainty in Hybrid Conflict: A Novel Approach and Model for Uncertainty Quantification in Intelligence Analysis," (2019).

10 Esther Gal-Or and Anindya Ghose, "The economic incentives for sharing security information," *Information Systems Research* 16, no. 2 (2005).

11 Loch K Johnson, *The Oxford handbook of national security intelligence* (Oxford University Press, 2010).

12 Adriana N Seagle, "Intelligence sharing practices within NATO: An english school perspective," *International Journal of Intelligence and CounterIntelligence* 28, no. 3 (2015).

13 F Korkisch, "NATO gets better intelligence," *IAS Reader, Strategy Paper* (2010).

Table 3. Outline of the Admiralty Code or NATO System

Source Reliability	Description
A – Completely reliable	No doubt of authenticity, trustworthiness, or competency; has a history of complete reliability
B – Usually reliable	Minor doubt about authenticity, trustworthiness, or competency; has a history of valid information most of the time
C – Fairly reliable	Doubt of authenticity, trustworthiness, or competency but has provided valid information in the past
D – Not usually reliable	Significant doubt about authenticity, trustworthiness, or competency but has provided valid information in the past
E – Unreliable	Lacking in authenticity, trustworthiness, and competency; history of invalid information
F – Reliability cannot be judged	No basis exists for evaluating the reliability of the source

Information Credibility	Description
1 – Completely credible	Logical, consistent with other relevant information, confirmed by independent sources
2 – Probably true	Logical, consistent with other relevant information, not confirmed
3 – Possibly true	Reasonably logical, agrees with some relevant information, not confirmed
4 – Doubtful	Not logical but possible, no other information on the subject, not confirmed
5 – Improbable	Not logical, contradicted by other relevant information
6 – Truth cannot be judged	The validity of the information cannot be determined

These characteristics of the NATO system on source credibility all relate to the notion of trust/trustworthiness. Trust is the attitude that takes to the trustworthiness of a source.[14] 'Trust is of central importance because quality is a perceived property and, thus, assessing the quality of an information source is essentially a matter of establishing to what extent one is willing to place trust in it'. [15]

Evaluating source intention as part of classification of information can be considered of vital importance.

In available literature about trust another property of trust is deemed important, besides the perceived competence of the source the perceived intent or agency of the source is essential for the trustworthiness of the source. [16] [17] Sources of information may have deviating goals, intentions and incentives that can alter their trustworthiness. Even though sources might be considered competent, their information might be comprehensive, consistent, accurate and up to date, they still may be suspected of following an agenda that is not in line with the receiver of information. [4]

While the competence of a source is often stable over time or might show gradual changes, intentions of sources, on the other hand, can be very volatile and might even change over-night (for example due to bribery, extorsion or other external pressure). Evaluating source intention as part of classification of information can be considered of vital importance. In the original NATO code, source intention might be considered a component of source reliability and assessed together with competence. Due to the specific importance of intent in the literature on trust and trustworthiness and the volatile character of source intention, in this study a separate assessments of source intention is proposed. In addition to the NATO code, a new classification scale is set up and tested in a practitioners panel (see Table 4).

14 Lea Viljanen, "Towards an ontology of trust" (paper presented at the International conference on trust, privacy and security in digital business, 2005).

15 Morten Hertzum et al., "Trust in information sources: seeking information from people, documents, and virtual agents," *Interacting with computers* 14, no. 5 (2002).

16 Katherine Hawley, *Trust: A very short introduction* (OUP Oxford, 2012).

17 Kieron O'Hara, "A general definition of trust," (2012).

Table 4. Proposed addition to the NATO code for classification of source intention

Source Intention	Description
I – Completely shared intentions	No doubt of source intention or aspiration, goals and objectives are in line; has a history of shared intentions
II – Usually shared intentions	Minor doubt about source intention or aspiration, goals and objectives are in line; has a history of shared intentions most of the time
III – Fairly shared intentions	Doubt of source intention or aspiration, goals and objectives might be in line; had shared intentions in the past
IV – Not usually shared intentions	Significant doubt about source intention or aspiration, goals and objectives might not be in line; had shared intentions in the past
V – No shared intentions	Lacking in transparency of source intention; goals and objectives might not be in line; had different intentions in the past
VI – Intention cannot be judged	No basis exists for evaluating the intention of the source

To explore the perceived trustworthiness and application of various information sources of security risk information, practitioners from the security domain are consulted. Different groups of practitioners participated in 1) a small brainstorming session to collect the most prominent possible sources of information, 2) a panel consultation to rank the source quality, and 3) a large-scale survey amongst security professionals to explore the application of these sources of information.

First a list of possible sources of risk information is composed during a brainstorming session with senior security professionals. This predefined list of possible sources of security risk information consists of 17 predefined sources as presented in the first column of Table 5.

For the ranking of the quality of these sources a practitioners panel is consulted. This panel consisted of 18 experienced security practitioners: on average 28 years of security experience, 83% followed specific security trainings, education level: associate degree 11%, bachelor degree 22%, master/PhD degree 67%. In an online consultation, the members of this panel are asked to rate the source reliability, information credibility and source intention of each of the predefined sources. The results of this consultation are analysed using a method for analysing Multiple-Criteria Decision Making: Fuzzy Technique for Order Performance by Similarity to Ideal Solution (FTOPSIS). The ranking is presented in table 5.

Table 5. Results of the FTOPSIS analysis, total results over the three criteria combined, in rank order followed by the results of each of the individual criteria: source reliability, information credibility, and source intention

Predefined information sources:	Total	Source Reliability	Information Credibility	Source Intention
	Ranking	Ranking	Ranking	Ranking
Experts	**1**	1	1	4
Personal experience	**2**	9	3	1
Science/scientific publications	**3**	3	1	7
Internal intelligence	**4**	4	5	5
External intelligence (government)	**5**	2	4	9
Peers	**6**	5	8	3
Personal training/education	**7**	8	7	2
Expert communities	**8**	6	6	6
Government or government agencies	**9**	7	9	10
Colleagues	**10**	12	12	7
External intelligence (commercial)	**11**	10	10	14
Consultants/consulting organisations	**12**	12	11	13
My 'Gut feeling'	**13**	11	13	12
Higher management	**14**	15	15	11
Supplier organisation s	**15**	16	14	15
Social media sources	**16**	14	16	16
Public sources like media	**17**	17	17	17

These results seem to confirm previous work in other domains that risk communication by government and industry is considered less trustworthy. [18] [19] [20] [21] Commercial sources like external commercial intelligence (11), consultants (12), and supplier organisations (15) are at the lower end of this ranking. They might contain too much marketing and are, therefore, considered less trustworthy.

Government sources rank somewhat higher: external government intelligence (5), government/government agencies (9). As other scholars concluded this lower perceived trustworthiness is primarily caused by deviating goals of both commercial and government risk information sources. The commercial and government sources indeed rank even lower on the source intention scale (last column of Table 5): commercial intelligence (14), consultants (13), supplier organisations (15), external government intelligence (9), government/government agencies (10).

In the main survey of this study 174 security professionals answered the research question: 'on what information source do you base your security risk assessment?' (see Table 6).

18 June Fessenden-Raden, Janet M Fitchen, and Jenifer S Heath, "Providing risk information in communities: Factors influencing what is heard and accepted," *Science, Technology, & Human Values* 12, no. 3/4 (1987).

19 David B McCallum, Sharon Lee Hammond, and Vincent T Covello, "Communicating about environmental risks: How the public uses and perceives information sources," *Health Education Quarterly* 18, no. 3 (1991).

20 Paul Slovic, James H Flynn, and Mark Layman, "Perceived risk, trust, and the politics of nuclear waste," *Science* 254, no. 5038 (1991).

21 Craig W Trumbo and Katherine A McComas, "The function of credibility in information processing for risk perception," *Risk Analysis: An International Journal* 23, no. 2 (2003).

Table 6. On what information source do you base your security risk assessment? Total results of the main survey, followed by the results of the practitioners panel (identical to Table 5)

Predefined information sources:	Application	Quality
	Ranking	Ranking
Experts	**1**	1
Personal experience	**2**	2
Internal intelligence	**3**	4
Peers	**4**	6
Personal training/education	**5**	7
Expert communities	**6**	8
External intelligence (government)	**7**	5
Government or government agencies	**8**	9
Science/scientific publications	**9**	3
Colleagues	**10**	10
External intelligence (commercial)	**11**	11
My 'Gut feeling'	**12**	13
Consultants/consulting organisation s	**13**	12
Public sources like media	**14**	17
Higher management	**15**	14
Supplier organisation s	**16**	15
Social media sources	**17**	16

It seems this new criterion as proposed in this paper, as an addition to the two criteria of the renowned NATO system, is of added value when evaluating the quality of sources of information.

The two rankings are, besides a few minor differences, similar. This indicates that the perceived high quality information sources, as assessed by the practitioners panel, are applied and perceived as important for risk assessments in praxis, as indicated by the group of respondents. The most remarkable difference between the rankings is the source: science/scientific publications. It is perceived a high-quality source (rank 3 by the panel) but seems to be less applied in daily praxis (rank 9 by the respondents). This might be explained by the additional proposed information quality criterion: source intention. The panellists assign a high source reliability to science/scientific publications (rank 3), the highest information credibility (rank 1 *ex aequo* with experts) but on source intention it is ranked at position 7. This means that there is at least some doubt on source intention or aspiration, goals and objectives might be in line (but this is not certain). Without the proposed additional criterion on information quality: source intention, this could not properly be explained. It seems this new criterion as proposed in this paper, as an addition to the two criteria of the renowned NATO system, is of added value when evaluating the quality of sources of information.

Individual characteristics of the respondents do not seem to be of much influence on the application of information sources during their security risk assessments. It seems that more experienced practitioners have more confidence in their own perception and less in government, commercial and social media sources.

Topic 2: perception, human processing of information

This section also presents brief summaries of two studies. In the first study the vulnerability of trained security professionals to known psychological and behaviour biases is examined. Does professional experience and training reduce the vulnerability to known and systematic distortion of judgment? This study is published in a full paper: "Biases in Security Risk Management: Do Security Professionals follow Prospect Theory in their Decisions?" in the *Journal of Integrated Security and Safety Science*.

The second study shows the results of several realistic security risk assessments. In these scenarios the descriptions are varied to explore the influence of more or less information. These experiments are based on the conjunction fallacy, predicting that likelihood estimates increase when case descriptions have more specific information, whereas they should actually decrease. This study is to be published in a full paper: "Bias and Noise in Security Risk Assessments, an Empirical Study on the Information Position and Confidence of Security Professionals," in : *Security Journal.*

Humans were found to follow decision behaviour that might be qualified as non-rational in the traditional economic sense.

The first study addresses the main research question: Are security professionals vulnerable to decision making biases as presented in Prospect Theory (PT)? PT has evolved in the 1970s and was driven by Amos Tversky and Daniel Kahneman, [22] who later received a Nobel prize for this work. After multiple experiments they concluded that the majority of people do not follow maximising theories in their decision making. The normative phenomenon of 'Homo Economicus', humans always choose the option with the highest perceived utility, was denied. Instead, humans were found to follow decision behaviour that might be qualified as non-rational in the traditional economic sense. These renown scholars identified multiple biases, which are systematic deviations from a norm or rational judgment.

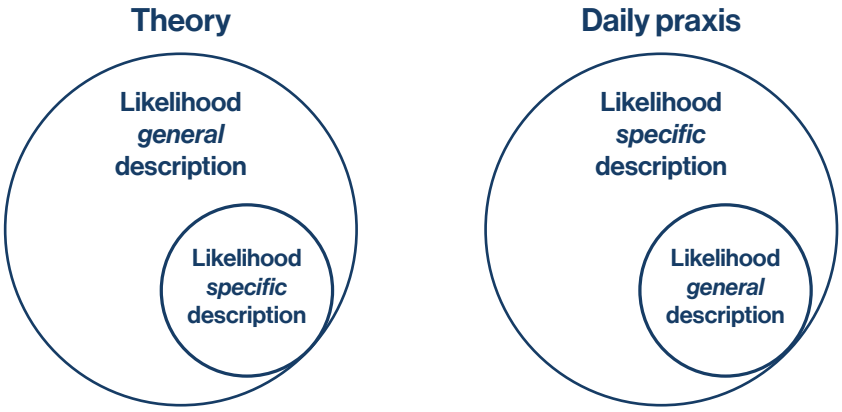
The original PT study focusses on decisions with two predefined options. The original experiments are, however, defined in financial loss and gain. This might not be representing security decisions. Therefore, in the second part of this study, the decision alternatives are redefined in security risk mitigation or reduction. The expectation is that security professionals, by the nature of their work and expertise, and confronted with limited, predefined, and given probabilities, could be less biased than lay people.

The results of this study clearly indicate that this expectation needs to be rejected. Based on the analysed results the vulnerability of security professionals to decision making biases using the original monetary gain and loss decisions showed an equal vulnerability to biases as lay people. Reformulating these experiments to reflect real life security decisions hardly changed the outcomes. The influence of the certainty effect, the non-linear preferences, the reflection effect, the lottery and insurance effect and the isolation effect on decision making by the majority of the sample of security professionals is clearly observed. This vulnerability to decision biases is revealed, on average, in decision behaviour of 70% of the sample of security professionals. In this short summary we will not further explain these biases, but the bottom line is that they all influence decision making in a way that the outcome is not maximised. The work of security professionals can be considered to be managing/mitigating risks, this study shows decision making that does maximise the outcome. It is safe to conclude that the studied biases can negatively affect optimal risk reduction.

22 Kahneman and Tversky, "Prospect theory: An analysis of decision under risk."

The core of the second study consists of cases testing the conjunction fallacy. The conjunction fallacy is a bias which identifies a flaw in logic reasoning. In theory the more specific a situation is, the less likely it would be compared to a less specific situation. The specific situation represents a subset of the generic situation. In practice, however, more details enhance the likelihood perception of humans (see Figure 1). In this study this phenomenon is explored to identify if it would influence the likelihood perception of professionals in real life risk assessments.

Figure 1. The conjunction fallacy explained



The case presented in this summary is a replication of the original problem statement as used by Kahneman and Tversky. [23] The context is reformulated to fit the security domain. As this reformulated problem shows the conjunction fallacy in plain sight, logic reasoning or recognition of the fallacy might influence the assessment of the respondents.

The reformulated problem consists of a short case description followed by a choice between two options. The respondents are asked to indicate which option they consider more likely. The first option has a general and short formulation. The second one is identical to the first, but is extended with more detailed information. Showing the two answers at the same time, in other words showing the conjunction rule, should or could guide the respondents to choose the shorter, more general, option. The second, more detailed, option, obviously is a sub-set of the first and should therefore be considered less likely.

23 Amos Tversky and Daniel Kahneman, "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment," *Preference, belief, and similarity: Selected writings by Amos Tversky* (2004).

This reformulated problem is presented to professionals of both the physical and cybersecurity domain:

Case introduction

Your organisation is a large, international, pharmaceutical corporation based in the EU. Your R&D department has focused the last months on research in developing a COVID-19 vaccine. This department made considerable progress and is considered to be one of the global front runners and ahead of other research institutes. Last week you discovered a serious attempt to steal information.

What is more likely:

- This attack is launched by an organised crime organisation
- This attack is launched by an organised crime organisation targeting IP (Intellectual Property) related to COVID-19 research

Note: this case is developed and presented to the respondents before - COVID-19 vaccines were available. At the time the surveys were conducted in both the physical and cybersecurity domain several pharmaceutical corporations around the world were in the race of developing vaccines and there were indications (in the press) of attempts of IP theft at these kinds of corporations. This case description can, therefore, be considered realistic.

Following the fallacy, retrieving more specific, detailed and recognisable information may lead the individual professional to consider a case, incident or threat more likely which in turn might lead to distorted risk assessments in organisations and society.

A total of 165 respondents answered the reformulated problem. 25.5% considered the first (short) option more likely, 74.5% the second (extended) one. In the physical security domain 58.8% of the respondents followed the fallacy and chose the extended option. Of the respondents active in the cybersecurity domain even 81.6% selected the extended option. The top threat in the cybersecurity domain in 2020 was IP theft by various threat vectors [24], while in the physical security domain the top threat in 2020 was malicious physical access [25]. The respondents originating from the cybersecurity domain, therefore, might relate more to the extended option. It fits their frame of reference, might lead to a stronger representation, recognition, emotion and thus availability. According to the theory of hints and the study of Brachinger and Monney [26] this explains the fall for the conjunction fallacy. An important consequence of this conclusion can be that professionals with domain expertise, and thus, a deeper subjective interpretation of so-called simple hints, and readily available information or even experience, [27] assess a higher likelihood to risks in their domain than non-domain experts.

In agreement with the hypothesis the results of this study clearly show the influence of the conjunction fallacy on the judgement of security professionals. As a consequence, security risk assessments by practitioners are probably influenced considerably by more detailed information. Following the fallacy, retrieving more specific, detailed and recognisable information may lead the individual professional to consider a case, incident or threat more likely which in turn might lead to distorted risk assessments in organisations and society. These findings have important implications for the professional security community and anyone depending on it.

24 "Top Cyberattacks of 2020 and How to Build Cyberresiliency," ISACA, 2020.

25 "Physical manipulation, damage, theft, loss, ENISA Threat Landscape," ENISA, 2020.

26 Hans Wolfgang Brachinger and Paul–André Monney, "The conjunction fallacy: explanations of the linda problem by the theory of hints," *International journal of intelligent systems* 18, no. 1 (2003).

27 Randy E Dumm et al., "The representative heuristic and catastrophe-related risk behaviors," *Journal of Risk and Uncertainty* 60, no. 2 (2020).

Overall conclusions

Our results make clear that professional security risk decision makers are as vulnerable to biases as lay people. This may lead to misconception of real-world risks. In roughly half of the real-life assessments detailed security risk information is lacking. This might even be considered overestimated (professionals perceive they have more detailed information than can be expected given the fact that the future is unpredictable by nature). Even if they are right, in half of the situations they seem to base their judgement on security risks without adequate information. The professionals indicate they can estimate a risk most of the time and even one in four assures they can always assess a security risk (even without information).

Overall, the security professionals show (over)confidence in their judgements, again even if they are aware there is no detailed evidence for their judgement. Our studies show that more experience leads to more (over)confidence and less need for more information. In other words: more experienced practitioners will base their judgment on less information. Even if they know information is lacking, they will decide without trying to retrieve more information.

Overall, the security professionals show (over)confidence in their judgements, again even if they are aware there is no detailed evidence for their judgement.

The sources of information with the highest perceived quality seem to be applied most in real-life praxis. The top 5 sources of information for security risks assessments, as applied by professionals, contain two individual sources: personal experience and training/education. These are considered very important and are in line with the previous conclusion that experienced practitioners are prone to use their expertise for their judgment. The other 3 in the top 5: experts, internal intelligence and peers, can be considered a part of the direct network of professionals. To influence them means these sources need to be influenced. There is also the danger of the resonation of information in a so called 'echo chamber' or bubble.

As our studies proved the vulnerability of professionals for well-known biases and heuristics, these might be leveraged to influence decisions and behaviour. Especially interesting is the conjunction fallacy: more detailed information raises the assessment of likelihood. If detailed information is communicated in the individual network of professionals, and resonated in echo chambers, it will most likely raise the likelihood perception.

The cornerstone of influence seems to be storytelling with details that trigger recognition of the individual risk assessor. More experience leads to more recognition which in turn might lead to a raised perception of likelihood.

Even the most experienced and best educated professional is human, and thus biased. The professional that is aware of this can reduce his/her own biased perception, but can also use it to influence others.

As we started this paper with McChrystal [28] we will also end with him:

'At the end of the day, we can't choose to have or have not biases – we have them. So we must identify and carefully consider them.'

28 McChrystal, *Risk, a user's guide*.

Bibliography

Brachinger, Hans Wolfgang, and Paul–André Monney. "The Conjunction Fallacy: Explanations of the Linda Problem by the Theory of Hints." *International journal of intelligent systems* 18, no. 1 (2003): 75-91.

Dumm, Randy E, David L Eckles, Charles Nyce, and Jacqueline Volkman-Wise. "The Representative Heuristic and Catastrophe-Related Risk Behaviors." *Journal of Risk and Uncertainty* 60, no. 2 (2020): 157-85.

"Physical Manipulation, Damage, Theft, Loss, Enisa Threat Landscape." ENISA, 2020.

Fessenden-Raden, June, Janet M Fitchen, and Jenifer S Heath. "Providing Risk Information in Communities: Factors Influencing What Is Heard and Accepted." *Science, Technology, & Human Values* 12, no. 3/4 (1987): 94-101.

Gal-Or, Esther, and Anindya Ghose. "The Economic Incentives for Sharing Security Information." *Information Systems Research* 16, no. 2 (2005): 186-208.

Gigerenzer, Gerd. *Risk Savvy: How to Make Good Decisions.* Penguin, 2015.

Hawley, Katherine. *Trust: A Very Short Introduction.* OUP Oxford, 2012.

Hertzum, Morten, Hans HK Andersen, Verner Andersen, and Camilla B Hansen. "Trust in Information Sources: Seeking Information from People, Documents, and Virtual Agents." *Interacting with computers* 14, no. 5 (2002): 575-99.

"Top Cyberattacks of 2020 and How to Build Cyberresiliency." ISACA, 2020.

Johnson, Loch K. *The Oxford Handbook of National Security Intelligence.* Oxford University Press, 2010.

Kahneman, Daniel, Olivier Sibony, and Cass R. Sunstein. *Noise, a Flaw in Human Judgment.* London: William Collins, 2021.

Kahneman, Daniel, and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica: Journal of the econometric society* (1979): 263-91.

Korkisch, F. "Nato Gets Better Intelligence." *IAS Reader, Strategy Paper* (2010): 1-2010.

McCallum, David B, Sharon Lee Hammond, and Vincent T Covello. "Communicating About Environmental Risks: How the Public Uses and Perceives Information Sources." *Health Education Quarterly* 18, no. 3 (1991): 349-61.

McChrystal, Stanley A. *Risk, a User's Guide.* New York: Penguin business, 2021.

Möller, Niklas. "The Concepts of Risk and Safety." Chap. 3 In *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, 55-85: Springer, 2012.

O'Hara, Kieron. "A General Definition of Trust." (2012).

Powell, Thomas, Serena Oggero, Joris Schook, and Emma Westerveld. "Dealing with Uncertainty in Hybrid Conflict: A Novel Approach and Model for Uncertainty Quantification in Intelligence Analysis." (2019).

Redmiles, Elissa M, Sean Kross, and Michelle L Mazurek. "How I Learned to Be Secure: A Census-Representative Survey of Security Advice Sources and Behavior." Paper presented at the Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016.

Seagle, Adriana N. "Intelligence Sharing Practices within Nato: An English School Perspective." *International Journal of Intelligence and CounterIntelligence* 28, no. 3 (2015): 557-77.

Simon, Herbert Alexander. *Models of Bounded Rationality: Empirically Grounded Economic Reason.* Vol. 3: MIT press, 1982.

Slovic, Paul, James H Flynn, and Mark Layman. "Perceived Risk, Trust, and the Politics of Nuclear Waste." *Science* 254, no. 5038 (1991): 1603-07.

Sunstein, Cass R. *Laws of Fear: Beyond the Precautionary Principle.* Cambridge University Press, 2005.

Trumbo, Craig W, and Katherine A McComas. "The Function of Credibility in Information Processing for Risk Perception." *Risk Analysis: An International Journal* 23, no. 2 (2003): 343-53.

Tversky, Amos, and Daniel Kahneman. "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment." *Preference, belief, and similarity: Selected writings by Amos Tversky* (2004): 221-56.

Viljanen, Lea. "Towards an Ontology of Trust." Paper presented at the International conference on trust, privacy and security in digital business, 2005.

Wright, George, and Peter Ayton. "Subjective Confidence in Forecasts: A Response to Fischhoff and Macgregor." *Journal of Forecasting* 5, no. 2 (1986): 117-23.

Johan de Wit owns a master's degree in Security Science and Management from Technical University Delft. Currently he holds a PhD research position at the faculty Technology, Policy and Management in the Safety and Security science group. He is exploring the characteristics of security risk assessments by security professionals. He is participating in scientific conferences and gives guest lectures. He is a member of various government committees, advisory boards, workgroups and communities of practice of norm institutions, government, academia and business associations in the Netherlands. Besides his research position at TU Delft Johan is working for Siemens Smart Infrastructure as Technical Officer Enterprise Security analyzing trends and developments in the (cyber) security domain. He is involved in global portfolio development and also supports customers and relations of Siemens in risk assessments and implementation of controls. He is a regular speaker and moderator at (international) conferences and seminars.

The Hague Centre
for Strategic Studies

HCSS
Lange Voorhout 1
2514 EA Hague

Follow us on social media:
@hcssnl

The Hague Centre for Strategic Studies
Email: info@hcss.nl
Website: www.hcss.nl