



LUND UNIVERSITY

How can the usefulness of capability assessments be improved?

Lindbom, Hanna; Tehler, Henrik; Frykmer, Tove; Uhr, Christian

2015

Document Version:

Peer reviewed version (aka post-print)

[Link to publication](#)

Citation for published version (APA):

Lindbom, H., Tehler, H., Frykmer, T., & Uhr, C. (2015). *How can the usefulness of capability assessments be improved?*. Paper presented at ESREL 2015.

Total number of authors:

4

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

How can the usefulness of capability assessments be improved?

H. Lindbom, H. Tehler, T. Frykmer

Division of Risk Management and Societal Safety, Centre for Societal Resilience, Lund University Centre for Risk Assessment and Management, Lund University, Lund, Sweden

C. Uhr

Division of Risk Management and Societal Safety, Centre for Societal Resilience, Lund University Centre for Risk Assessment and Management, Lund University, Lund, Sweden

Swedish Civil Contingencies Agency, Sweden

ABSTRACT: Capability assessments are used in disaster risk management to facilitate decision making regarding capability increasing measures. Recent studies suggest that two factors might determine their usefulness for decision making: whether they include descriptions of an actor's available *resources*, and how well the actor can accomplish a specific *task* in case of a disaster. The experimental study presented here aimed at investigating the extent to which these factors influence the perceived usefulness for decision making of four hypothetical capability assessments, differing in whether they contain information regarding the two factors. 89 participants from the Swedish fire and rescue services were randomly assigned to rate the perceived usefulness of one of the four versions and the results show that the presence of both factors makes assessments more useful, and resources are especially important to include.

1 INTRODUCTION

Capability assessments and capabilities-based planning are becoming increasingly popular as part of the preparedness for disasters in order to prepare for a wide variety of risks and threats instead of specific scenarios (Programme National Security 2007). Several countries, including Australia, The Netherlands, New Zealand, Sweden, The UK and The USA use capability assessments as part of their emergency preparedness (Australian Capital Territory 2012, Cabinet Office 2014, Dutch Ministry of Interior and Kingdom Relations 2009, Homeland Security 2013, Houdijk 2010, Ministry of Civil Defence and Emergency Management n.d., Swedish Civil Contingencies Agency (MSB) 2014). The purpose of the assessments, similar to the purpose of risk assessments, is often to facilitate decision making in order to increase capability (Abt et al. 2010, Bier 2001a, Johansen & Rausand 2014, Palmqvist et al. 2014). Capability is often defined as resources (Lindbom et al. 2015) and existing capability assessment methods often specify indicators for capability, e.g. if plans exist and if drills have been performed. Sometimes the indicators are weighed together into a final capability index (Palmqvist et al. 2014).

However, recent studies suggest that the conceptual as well as the methodological basis for capability assessments are unclear (Lindbom et al. 2015, Palmqvist et al. 2014). Lindbom et al. (2015) also report on a study of the Swedish disaster risk management system which indicates that two important factors that might determine the usefulness of a capability assessment are (1) the extent to which the capability assessments include descriptions of the *resources* available for dealing with a disruptive event, and (2) the extent to which they include descriptions of how well the actor in question (e.g. a specific fire and rescue service) is judged to be able to accomplish a specific *task* in the case of a disruptive event. The aim of the experimental study presented in this paper was to investigate the extent to which the two factors *resources* and *task* influence the perceived usefulness for decision making of hypothetical capability assessments.

Following this introduction, the outline of the paper is as follows. First we report on previous research on communication of risk and capability assessments to decision makers. We then present the experiment, including participants, procedure, analysis and results. That is followed by a discussion about the findings and a conclusion.

2 COMMUNICATION TO DECISION MAKERS

How to communicate results from risk and capability assessments is important since these assessments should form the basis for making decisions about how to increase capability and reduce risk. As far as we know, there is limited previous research focusing on using capability assessments as a basis for decisions. Since the purpose of both capability assessments and risk assessments is to facilitate decision making and the fact that capability and risk are closely related (Lindbom et al. 2015), one could use research on risk assessments as decision support as a point of departure and draw parallels to capability assessment. However, research on how to communicate risk assessments to decision makers is also limited (Bier 2001a) and below, we will summarize what we have found on how risk assessments in general should be structured in order to facilitate decision making.

An important note to start with is that the best presentation format may vary depending on the disciplinary background of the decision maker, the nature of their concerns and how the risk assessment will be used (Arvai 2007, Bier 2001a, Comes et al. 2012, Doyle et al. 2014, Fraser-Mackenzie et al. 2014, Kristensen et al. 2006, Ohanian et al. 1997, Williams & Paustenbach 2002). Also, the risk perspectives of the persons taking part in risk communication, i.e. how they understand concepts such as risk, uncertainty and probability, affect what is communicated and how it is perceived (Veland & Aven 2013).

Researchers agree that both qualitative and quantitative descriptions should be available for the decision maker (Aven & Renn 2009, Bier 2001a, Williams & Paustenbach 2002) and that uncertainties should be discussed (Aven & Renn 2009, Bier 2001a, Ohanian et al. 1997, Rother 2014, Williams & Paustenbach 2002). Assumptions and their implications should be presented (Davis & Kahan 2006, Ohanian et al. 1997, Williams & Paustenbach 2002), as well as alternatives and rationales for action (Ohanian et al. 1997, Rother 2014) and a comparison to other risks (Ohanian et al. 1997). The decision basis should be a summary of the whole risk assessment (Davis & Kahan 2006, Williams & Paustenbach 2002) and be comprehensive, credible (Bier 2001a), understandable (Bier 2001a, Rother 2014), logical (Davis & Kahan 2006) and uncomplicated (Rother 2014).

As opposed to just facilitating the decision, what does exist is plenty of research on how decisions are made in general and under uncertainty in particular, and whether those decisions were correct or not. Central is the seminal work by Tversky & Kahneman

(1974) on how heuristics and biases affect decision making. Following that work, much research has been carried out on how uncertainties are interpreted and how they should be expressed. Just to mention a few, Dieckmann et al. (2012) found that a combination of numerical values and verbal labels as a means to express uncertainties is to prefer since this will allow the decision makers to choose the type of expression they prefer. Further, Mauro & Slovic (2010) conclude that decision makers are less willing to make decisions when ambiguity is explicit, for example through the use of second order probabilities, and Doyle et al. (2014) studied how verbal probability expressions are transferred into numerical equivalents.

Whereas how to express uncertainties has been widely studied, not much has been done on how to express events and/or scenarios (Fischhoff 1994) and the importance of scenarios is advocated by those in favour of scenario-based decision making where descriptions of alternative future realities are created in the form of causal story-like scenarios against which strategies can be tested (Harries 2003).

Another well-studied area when it comes to risk communication is on communication to the public and some researchers say that risk communication to decision makers is similar to communicating risk to the public (Bier 2001a, Frank 2002). Bier (2001b), in her compilation of research on communication to the public, says that much research has been carried out on the format of the risk communication message (e.g. if verbal or numerical formats are preferred and what results various communication approaches yield, e.g. one *should* do..., or one *may want to* do...). She also summarises research on using risk comparison for communicating risk to the public, and how differences in the audience (gender, worldview, socioeconomic status) affect how a risk message is being perceived. The overview however concludes that there seems to be little systematic impact by these factors on the effectiveness of the resulting communication. Further, Bier reviews the research on stakeholder participation, credibility and trust and mental models. For the latter she reports that three barriers are common in terms of hindering understanding of the communicated message: lack of familiarity with a concept or term, lack of mental models relevant to the subject, and the existence of misconceptions. Her conclusion is that it is important to make an assessment of what the decision maker needs from the decision basis in order to be able to make a decision. Thus, maybe it is not sufficient to find *the* best general format since that might vary from situation to situation and from decision maker to decision maker.

To conclude this chapter, what researchers seem to agree on is that how the risk assessments should be used, and by whom, should influence how it is communicated (Arvai 2007, Comes et al. 2012, Fraser-Mackenzie et al. 2014, Williams & Paustenbach 2002), that how risk is communicated to decision makers is important (Dilla & Stone 1997, Kühberger 1995, Scurich & John 2011) and that small variations in how risk is presented will change the decision made by the decision maker (Dilla & Stone 1997, Kühberger 1995). Whereas the communication of uncertainties, and consequences to some extent, have previously been studied, what seems to be lacking in the scientific literature on communication of risk (and capability) to decision makers is how the communication format of task and resources (the two components suggested by Lindbom et al. 2015) affects the perceived usefulness of the decision basis and in turn the decisions made based on it. The experiment described in the next section is an initial attempt to rectify this lack of knowledge.

3 EXPERIMENT

3.1 *Participants*

Participants were directed to the online survey tool Fluidsurveys (FluidSurveys 2015) through a snowball approach via email contact with individuals in the Swedish fire and rescue services. In total, 89 participants completed the survey: 21 working with only operational tasks, 3 working only with preventive tasks, and 65 working with both. Two participants did not report where they worked, but the remaining 87 participants represent 28 Swedish fire and rescue services from a total of 160 (Swedish Association of Local Authorities and Regions 2015). The participants' work experience in the fire and rescue services ranged from less than 1 to over 30 years (average 16 years).

3.1 *Procedure*

Fluidsurveys randomly assigned each participant to one of four experimental groups. Random assignment is an experimental design to randomly allocate participants where each participant has the same probability ending up in any of the four groups (Cunningham & Wallraven 2011 p. 238) and this helps reducing the error variance (Bordens & Abbott 2010 p. 293).

We used a between-subject approach (Bordens & Abbott 2010, Cunningham & Wallraven 2011) and all participants were told that they were to see one of four fictive capability assessments for the fire and

rescue service's capability to extinguish forest fires. They were asked to assess how useful the capability assessment is to help them make decisions in order to increase the capability of the fire and rescue service to handle forest fires, for example through training or investing in new equipment. First all participants answered general questions about their age, what fire and rescue service they work for, how many years of experience they have from the fire and rescue services, how many times they have been at a forest fire, what they judge their knowledge about extinguishing forest fires to be, and if they work with only preventive tasks, operative tasks or both.

Thereafter they were shown one of the four capability assessments. Table 1 shows the information available in each version. We used a 2x2 between-subject design with the two factors task and resources and two levels (yes/no) for each factor.

Thus, the most limited in terms of information available was a capability assessment containing no information about available resources and no descriptions of tasks that can be performed (VERSION 1: RESOURCES = no, TASK = no). The most informative experimental condition contained information on both resources and task (VERSION 4: RESOURCES = yes, TASK = yes).

Based on the capability assessment shown to them, they were asked to rate four statements: the assessment is useful as a basis for decision regarding measures that can enhance the fire and rescue service's capability to extinguish forest fires; the assessment is easy to use as a basis for decision; the assessment is realistic, i.e. the example could come from a Swedish municipality; the assessment is clear and logic. The rating was done using a seven-level Likert item ranging from 'I completely agree' to 'I strongly disagree'. The option 'Don't know' was also available.

The participants were given the opportunity to provide free-text answers on how the example could be enhanced in order to become more useful as a basis for decision.

The task took 9.5 minutes to complete on average. Of the 89 persons included in the study, 29 were assigned to group 1, 19 to group 2, 18 to group 3 and 23 to group 4.

Table 1. The information available in each version of the hypothetical capability assessments used in the experiment

Version	Components
1	A
2	A+B+map(a+b)
3	A+C+map(b)
4	A+B+C+map(a+b)
Description of components	
A	A description of the municipality in terms of population and forest areas, a description of the group that assessed the capability and their final conclusion that the fire and rescue service's capability to handle a forest fire is good but with some deficiencies.
General info	
B	A description of the number of fire fighters in each of the six districts in the municipality, the type of forest and topography in four different zones of the forest and the number of houses within each zone. It is explicit that this information was used as point of departure in the assessment.
Resources	
C	A description of how fast the fire and rescue service can extinguish a fire in each of the four zones and how large the affected area will be. It is explicit that this information was used as point of departure in the assessment.
Task	
Map	a) Shows where the districts' fire stations are located, the number of fire fighters at each station. b) Shows the different zones.

3.2 Analysis method

Due to the use of ordinal scales we utilize non-parametric tests to test for statistical significance as these tests make no distributional assumptions about the population under investigation (Bordens & Abbott 2010). We consider each statement in turn and compare the rating provided by the four groups, two groups at the time. We do not compare groups 1 and 4, and 2 and 3, since this would mean changing levels for both factors concurrently. For the test we use a Wilcoxon rank sum test, which is a non-parametric equivalent to the t-test (Bordens & Abbott 2010, Cunningham & Wallraven 2011), to test the significance of the difference between two independent groups. The participants who answered 'Don't know' for the analysed theme were not included in the analysis of that specific theme. Therefore, the number of subjects included in the analysis, as reported in the tables below, varies.

For the Wilcoxon rank sum test we report the result of the test (W and Z) and the p -value, which is the probability that this statistic could occur by sampling error if the null hypothesis is true (Cunningham & Wallraven 2011). We assume that a result is statistically significant when $p < 0.05$ (two-tailed test). We also present the mean and median values for each statement and each group.

3.3 Results

Tables 2-5 present the results from the statistical analyses. Each table includes the average values for each group within each theme, as well as the results from the Wilcoxon rank sum test.

Table 2. Usefulness for decision making

Group	Mean	Median	Number of subjects
1	3.741	4	27
2	5.167	5.5	18
3	4.235	5	17
4	5.478	6	23
Comparison	p -value	Z	W
1 and 2	0.008285	-2.640	509.0
1 and 3	0.3511	-0.9325	570.0
2 and 4	0.8252	-0.2208	369.5
3 and 4	0.003320	-2.936	246.5

Table 3. Ease of use for decision making

Group	Mean	Median	Number of subjects
1	3.259	2	27
2	4.632	5	19
3	4.235	5	17
4	5.045	5	22
Comparison	p -value	Z	W
1 and 2	0.01463	-2.4414	527.0
1 and 3	0.06326	-1.857	532.5
2 and 4	0.5623	-0.5795	377.0
3 and 4	0.1769	-1.3503	293.5

Table 4. Realistic

Group	Mean	Median	Number of subjects
1	5.071	6	28
2	5.611	6	18
3	5.688	5.5	16
4	5.318	6	22
Comparison	<i>p</i> -value	Z	W
1 and 2	0.2115	-1.249	605.0
1 and 3	0.3130	-1.0090	590.0
2 and 4	0.4993	-0.6757	393.0
3 and 4	0.6539	0.4484	327.0

Table 5. Clear and logic

Group	Mean	Median	Number of subjects
1	3.321	3	28
2	5.111	6	18
3	4.706	5	17
4	5.217	5	23
Comparison	<i>p</i> -value	Z	W
1 and 2	0.001405	-3.194	518.5
1 and 3	0.01202	-2.5117	539.0
2 and 4	0.7727	0.2889	389.0
3 and 4	0.3517	-0.9313	315.5

4 DISCUSSION

The aim of the study presented was to investigate the extent to which the two factors resources and task influence the perceived usefulness for decision making of four versions of hypothetical capability assessments. Looking first at the averages presented in tables 2-5, the versions where resources and/or tasks are presented (versions 2, 3 and 4) are perceived more useful than when none of them are presented (version 1). Further, the versions in which resources are presented (2 and 4) are perceived more useful, easy to use and logic compared to those where resources are not presented (1 and 3). When it comes to how realistic the versions are perceived, the averages for the four versions are similar. Thus, both the factors *resources* and *tasks* affect the perceived usefulness of the capability assessments.

Taking also the Wilcoxon rank sum test into consideration, the results show that the difference is sig-

nificant when it comes to resources, but not with regards to task. When it comes to clarity and logical consistency, versions 2 and 3 are perceived significantly better than version 1.

The respondents were asked to provide free-text answers on how the capability assessment they had seen could be changed to better work as a basis for decision. 51 respondents chose to do this (17 for version 1, 11 for version 2, 11 for version 3 and 12 for version 4). In total, 20 of the respondents (version 1: 8; version 2: 3; version 3: 4; version 4: 5) comment about insufficient descriptions of resources. What they mention as lacking is available vehicles, material, collaboration partners and agreements, education, exercise drills, knowledge, and competence. Only two respondents mention insufficient descriptions of tasks (version 1: 1; version 2: 1). They want to receive information about what actually can be done, and what the shortages are. When it comes to the assessments' usefulness for decision making, 13 respondents comment on this and the most common comment is that there is too little detail in the assessment in order for it to be useful for decision making (7 comments; 4 comments for version 1, and one each for versions 2, 3 and 4). Others want more motivations related to how the conclusion 'the capability is high but with some deficiencies' was made (3 comments), information about what is prioritised in the assessment (1 comment) and information about who made the assessment (1 comment). One person comments that with the assessment a decision maker can decide that *something* needs to be done, but not *what* needs to be done (authors' emphasis) and one person says that the assessment should include more versions of the scenario, e.g. the most credible and the worst credible.

Several comments also relate to describing the scenario and the affected area (20 comments; 4 comments for version 1, 7 for version 2, 6 for version 3 and 3 for version 4). Useful information would be descriptions, in text or on a map, of vegetation type, humidity, wind, weather, time of the day, roads that can be used for reaching the area, where water can be found, buildings etc. Two respondents from group 1 (who did not receive a map) wanted a map. Only one respondent (from group 3) said that assessments of the consequences were lacking, in terms of 'is it ok to be able to extinguish the fire within 4 hours?'

When reading the comments, it also becomes clear that some respondents have interpreted the capability assessment as a basis for decision making to be used in the acute phase of a forest fire. This is clear in at least six comments (three from group 2, two from group 3 and one from group 4). All of the-

se respondents work both with operational and preventive tasks. It is possible that more respondents have answered the survey based on that the assessment was to be used in the acute phase of an emergency, whereas our intention was to study the use of decision basis in the preparedness phase. In future studies, it would be interesting to look at how the time aspect influences the perceived usefulness of the decision basis, i.e. what factors influence the perceived usefulness when decisions are to be made under time pressure vs. when there is no time pressure.

Other aspects of the present study that would be interesting to further investigate in future studies are how experience and knowledge of the area affect perceived usefulness. Based on the answers from the participants, there seems to be a difference in perceived usefulness between the versions with a positive correlation (Spearman rank correlation, $\tau=0.4047$) between experience and perceived usefulness for version 2 (RESOURCES=yes, TASK=no) and a negative correlation ($\tau=-0.2403$) between experience and perceived usefulness for version 3 (RESOURCES=no, TASK=yes). However, the findings are not significant ($p=0.09577$, $p=0.3529$), but what this could indicate is that experienced persons find resources alone sufficient because they can fill in the blanks themselves on how the resources can be used to perform a task and if they are sufficient for the scenario at hand. A novice on the other hand, may not have the knowledge to do this and needs the support of explaining how the resources can be used and for what.

The difference in the perceived usefulness between experienced and inexperienced persons is important when it comes to multi-actor capability assessments. In today's society when actors are becoming increasingly dependent on one another, their response and the result of their efforts depend on each other's response. The overall result depends on the actors' joint effort to respond to the event. Say for example that Actor A is a fire and rescue service and Actor B is a hospital. In case of a serious building fire, the ability to save the lives of people trapped in the burning building depends on the performance of both actors. The hospital treats the people exposed to heat and smoke when they arrive at the hospital. The impact, for example in terms of the number of fatalities, depends on the ability of the hospital to treat the patients, but also on the conditions of the patients when they arrive at the hospital. This in turn depends on the ability of the fire and rescue service to quickly rescue the people from the burning building. Therefore an assessment of capability would also need to take these dependencies into account. If the hospital assesses their capability to

treat people exposed to heat and smoke due to a building fire, it should be based on the fire and rescue service's capability assessment on how fast they can rescue people from the burning building. In a capability assessment based on resources only, the ability of the fire and rescue service would be expressed for example in terms of the number of fire fighters and fire trucks, and if there is a standard operating procedure for how to rescue people from burning buildings. However, this information in itself is not enough, or even relevant, for the hospital when they want to assess their ability to treat the victims. They probably do not have the domain specific competence to interpret what five fire trucks mean in terms of the condition of the patients when they arrive at the hospital. Instead it would probably be more valuable for the hospital to know the impact of the fire and rescue services work, in this case how long time the victims have been exposed to heat and smoke. This is supported by Bier (2001a p. 156) saying that when "communicating risk information to professionals in other disciplines, attention should be paid to ensure that the topics of greatest relevance to those disciplines are emphasized". So if one wishes to take dependencies into account when assessing capability, focusing on resources only comes with limitations when it comes to sharing information between actors.

A limitation of the present study is that the term *usefulness* has not been defined. Thus, we do not know what the respondents thought about when it came judging whether the capability assessment was *useful* for decision making. Possible interpretations could be that decisions can be made faster if the decision basis is more useful, or that the decisions will be better. A future study investigating how participants interpret the concept of usefulness would help clarify this.

Another note on usefulness is that we have studied the perceived usefulness of capability assessments for decision making. It should be stressed that the usefulness and the decision making is not the ultimate goal of these assessments. Rather, they are the means to increase capability and achieve the goal of reducing impact and likelihood of disasters. However, it is difficult to study various capability assessments' effect on actual losses, for example since disasters do not strike very often and the collection of data would therefore be difficult. Therefore, the approach taken in this study is more practical when investigating the issue. Our assumption then is that more useful assessments lead to better decisions, which in turn increase capability and reduce likelihood and impact of future disasters. In order to test this assumption, it would be interesting to in the fu-

ture study the actual decisions made based on various capability assessments.

5 CONCLUSION

We have conducted experiments to study the perceived usefulness of capability assessments for decision making. 89 participants from the Swedish rescue services were randomly assigned to rate the perceived usefulness for decision making of one of four versions of capability assessments. The versions differed in whether they contained information or not regarding resources that could be used, and tasks that could be performed, in the event of a forest fire.

Based on the results, we conclude that whether resources and/or tasks are described in the capability assessments affect the perceived usefulness for decision making. Thus, the usefulness of capability assessments for decision making can be improved by including the two components resources and task.

However, the experience of the decision maker within the area of interest also seems to be an important factor. Inexperienced decision makers (as opposed to experienced decision makers) seem to find tasks to be more important to include in the decision basis. Although the results in this study regarding the role of experience were not significant, it deserves attention in future studies. The reason for this is that the complexity of today's society calls for multi-actor capability assessments and one cannot assume that one actor has domain specific knowledge of other actors' area of expertise and therefore, only a description of available resources will not be enough.

ACKNOWLEDGEMENTS

We thank the Swedish Civil Contingencies Agency for funding the research on which the present paper is based.

REFERENCES

- Abt, E., Rodricks, J. V., Levy, J. I., Zeise, L., & Burke, T. A. 2010. Science and decisions: advancing risk assessment. *Risk Analysis*, 30(7): 1028–1036.
- Arvai, J. L. 2007. Rethinking of risk communication: lessons from the decision sciences. *Tree Genetics & Genomes*, 3(2): 173–185.
- Australian Capital Territory 2012. Emergencies Act 2004, republication no 20.
- Aven, T., & Renn, O. 2009. The role of quantitative risk assessments for characterizing risk and uncertainty and delineating appropriate risk management options, with special emphasis on terrorism risk. *Risk Analysis*, 29(4): 587–600.
- Bier, V. 2001a. On the state of the art: risk communication to decision-makers. *Reliability Engineering and System Safety*, 71(2): 151–157.
- Bier, V. 2001b. On the state of the art: risk communication to the public. *Reliability Engineering & System Safety*, 71(2): 139–150.
- Bordens, K. S., & Abbott, B. B. 2010. *Research design and methods: A process approach*. New York, NY: McGraw-Hill.
- Cabinet Office 2014. Preparation and planning for emergencies: the national resilience capabilities programme - detailed guidance. <http://www.cabinetoffice.gov.uk/content/capabilities-programme>. Retrieved March 30, 2015, from <http://www.cabinetoffice.gov.uk/content/capabilities-programme>
- Comes, T., Wijngaards, N., & Schultmann, F. 2012. Efficient Scenario Updating in Emergency Management. In Leon Rothkrantz, Jozef Ristvej, Zeno Franco (eds.) *Proceedings of the 9th International ISCRAM Conference, Vancouver, 22-25 April 2012*. Vancouver: Simon Fraser University.
- Cunningham, D. W., & Wallraven, C. 2011. *Experimental Design: From User Studies to Psychophysics [E-book]*. Hoboken: CRC Press.
- Davis, P. K., & Kahan, J. P. 2006. Theory and methods for supporting high-level decision making. In Dawn Trevisani (ed.) *Proc. SPIE 6227, Enabling Technologies for Simulation Science X, Kissimee, 17-18 April 2006*.
- Diekmann, N. F., Mauro, R., & Slovic, P. 2010. The Effects of Presenting Imprecise Probabilities in Intelligence Forecasts. *Risk Analysis*, 30(6): 987–1001.
- Diekmann, N. F., Peters, E., Gregory, R., & Tusler, M. 2012. Making sense of uncertainty: advantages and disadvantages of providing an evaluative structure. *Journal of Risk Research*, 15(7): 717–735.
- Dilla, W. N., & Stone, D. N. 1997. Representations as Decision Aids: The Asymmetric Effects of Words and Numbers on Auditors' Inherent Risk Judgments. *Decision Sciences*, 28(3): 709–743.
- Doyle, E. E. H., McClure, J., Johnston, D. M., & Paton, D. 2014. Communicating likelihoods and probabilities in forecasts of volcanic eruptions. *Journal of Volcanology and Geothermal Research*, 272: 1–15.
- Doyle, E. E. H., McClure, J., Paton, D., & Johnston, D. M. 2014. Uncertainty and decision making: Volcanic crisis scenarios. *International Journal of Disaster Risk Reduction*, 10: 75–101.
- Dutch Ministry of Interior and Kingdom Relations 2009. *Working with scenarios, risk assessment and capabilities in the National Safety and Security Strategy of the Netherlands*.
- Fischhoff, B. 1994. What forecasts (seem to) mean. *International Journal of Forecasting*, 10: 387–403.
- FluidSurveys. 2015. Online Survey Software. Retrieved March 9, 2015, from <http://fluidsurveys.com/>
- Frank, M. V. 2002. Principles of risk communication: A brief primer from the experience of a risk practitioner. In Bruce Battrick (ed.) *Proceedings of Joint ESA-NASA Space-Flight Safety Conference, Noordwijk, 11-14 June 2002*. Noordwijk: ESA Publications Division.

- Fraser-Mackenzie, P., Sung, M.-C., & Johnson, J. E. V. 2014. Toward an Understanding of the Influence of Cultural Background and Domain Experience on the Effects of Risk-Pricing Formats on Risk Perception. *Risk Analysis*, 34(10): 1846–1869.
- Harries, C. 2003. Correspondence to what? Coherence to what? What is good scenario-based decision making? *Technological Forecasting and Social Change*, 70: 797–817.
- Homeland Security 2013. *Capability Estimation. Comprehensive Preparedness Guide (CPG) XXX. Pre-decisional working draft. For review purposes only.*
- Houdijk, R. 2010. *Regional risk assessment in The Netherlands - an introduction.* The Hague.
- Johansen, I. L., & Rausand, M. 2014. Foundations and choice of risk metrics. *Safety Science*, 62: 386–399.
- Kristensen, V., Aven, T., & Ford, D. 2006. A new perspective on Renn and Klinke's approach to risk evaluation and management. *Reliability Engineering & System Safety*, 91(4): 421–432.
- Kühberger, A. 1995. The Framing of Decisions: A New Look at Old Problems. *Organizational Behaviour and Human Decision Processes*, 62(2): 230–240.
- Lindbom, H., Tehler, H., Eriksson, K., & Aven, T. 2015. The capability concept - on how to define and describe capability in relation to risk, vulnerability and resilience. *Reliability Engineering & System Safety*, 135: 45–54.
- Ministry of Civil Defence and Emergency Management n.d. National Capability Assessment. Retrieved March 11, 2015, from <http://www.civildefence.govt.nz/cdem-sector/monitoring-and-evaluation/national-capability-assessments/>
- Ohanian, E. V., Moore, J. A., Fowle III, J. R., Omenn, G. S., Lewis, S. C., Gray, G. M., & North, D. W. 1997. Risk characterization: A bridge to informed decision making. *Fundamental and Applied Toxicology*, 39(2): 81–88.
- Palmqvist, H., Tehler, H., & Shoaib, W. 2014. How is capability assessment related to risk assessment? Evaluating existing research and current application from a design science perspective. In *Proceedings of the Twelfth International Conference on Probabilistic Safety Assessment and Management (PSAM), Honolulu, June 22-27 2014.*
- Programme National Security 2007. *National security, strategy and work programme 2007-2008.* Breda: Programme national security, Ministry of the interior and kingdom relations.
- Rother, H.-A. 2014. Communicating pesticide neurotoxicity research findings and risks to decision-makers and the public. *NeuroToxicology*, 45: 327–337.
- Scurich, N., & John, R. S. 2011. The effect of framing actuarial risk probabilities on involuntary civil commitment decisions. *Law and Human Behavior*, 35(2): 83–91.
- Swedish Association of Local Authorities and Regions 2015. Map municipal fire and rescue services [Karta kommunal räddningstjänst]. Retrieved March 9, 2015, from <http://skl.se/download/18.5b06742214b3091389c5a949/1425044506667/Kommunal-raddningstjanst-karta-2015.pdf>
- Swedish Civil Contingencies Agency (MSB) (2014). National risk and capability assessment [Nationell risk- och förmågebedömning]. Retrieved March 11, 2015, from <https://www.msb.se/sv/Forebyggande/Krisberedskap/Nationell-risk-och-formagebedomning/>
- Tversky, A., & Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157): 1124–1131.
- Veland, H., & Aven, T. 2013. Risk communication in the light of different risk perspectives. *Reliability Engineering & System Safety*, 110: 34–40.
- Williams, P. R. D., & Paustenbach, D. J. 2002. Risk characterization: principles and practice. *Journal of Toxicology and Environmental Health. Part B, Critical Reviews*, 5(4): 337–406.