Defence Research and Development Canada | Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

Risk Assessment: Challenges and Recommendations

Kyungryun (Cathy) Pak

Lynne Genik

DRDC Centre for Security Science

Defence R&D Canada – Centre for Security Science

Technical Note
DRDC CSS TN 2012-032
December 2012

Canada

Principal Author

Kyungryun (Cathy) Pak

DRDC Centre for Security Science

Approved by

Dr. Denis Bergeron

DRDC Centre for Security Science
Manager - Decision Support

Approved for release by

Dr. Andrew Vallerand

DRDC Centre for Security Science
Document Review Panel Chairman

Abstract

This paper presents a literature review on risk assessment (RA) that focuses on some of the challenges that arise when conducting RAs. This work builds upon a previous literature search on RA, identifying several key references and synthesizing common issues and recommendations. An assessment of quantitative, qualitative, integrated and comparative RA approaches, and a discussion on availability, representativeness and anchor-and-adjustment psychological heuristics are included. The following recommendations are provided: that risk practitioners recognize and communicate uncertainty; adopt scientific practices such as documentation, validation, peer review, and publishing; and improve transparency. The literature review was undertaken as a part of the Defence Research and Development Canada (DRDC) and Emergency Management British Columbia (EMBC) collaborative project and is intended to be a resource for DRDC, EMBC, and other partners.

Résumé

Le présent document contient une analyse documentaire sur l'évaluation des risques, axée sur certains des défis qui se présentent lors des évaluations des risques. Ce travail s'appuie sur une recherche documentaire antérieure sur l'évaluation des risques qui a permis de repérer plusieurs documents de référence clés et de faire une synthèse des questions et des recommandations courantes. Ce document comprend une évaluation des méthodes quantitatives, qualitatives, intégrées et comparatives d'évaluation des risques ainsi qu'une discussion sur les heuristiques psychologiques (heuristique de disponibilité, de représentativité et d'ancrage et d'ajustement). Les recommandations suivantes ont été formulées : que les spécialistes de la gestion du risque reconnaissent et communiquent l'incertitude, qu'ils adoptent des pratiques scientifiques comme la documentation, la validation, l'évaluation par les pairs et la publication et qu'ils améliorent la transparence. L'analyse documentaire a été entreprise dans le cadre d'un projet de collaboration entre Recherche et développement pour la défense Canada (RDDC) et Emergency Management British Columbia (EMBC). Il s'agit d'un outil que RDDC, EMBC et d'autres partenaires peuvent utiliser comme ressource.

Executive summary

Risk Assessment: Challenges and Recommendations

Kyungryun (Cathy) Pak, Lynne Genik, DRDC Centre for Security Science; DRDC CSS TN 2012-032; Defence R&D Canada – CSS; December 2012.

Background: As a part of the collaborative project between Defence Research and Development Canada (DRDC) and Emergency Management British Columbia (EMBC), literature searches were documented in the areas of risk assessment (RA) and critical infrastructure (CI). The literature search documents consist of approximately 200 references for each of the two subject areas. Following the completion of these documents, DRDC decided to build on its work by conducting a literature review for RA, with the aim of identifying common challenges and limitations of RA practices along with recommendations. This paper presents the results of this literature review.

Method: The literature review was conducted by drawing upon the references included in DRDC's literature search document for RA. A limited amount of additional research was also conducted, primarily through the use of online databases such as the DRDC research database and the University of Ottawa library database. The principal author selected several references that provided discussions on the concepts, limitations, and recommendations concerning RA. Key elements of these discussions were identified and synthesized in order to provide an overview of important considerations pertaining to RA.

Results: A synthesis of the existing literature on RA is presented in this paper. Firstly, the quantitative, qualitative, integrated, and comparative approaches to RA are described, with a focus on highlighting the limitations of these methods. The paper then discusses the implications of psychological heuristics on the human ability to make accurate judgements. These include the availability heuristic; the representativeness heuristic; and the anchor-and-adjustment heuristic. Lastly, recommendations are provided, and risk practitioners are advised to identify and communicate uncertainty; adopt scientific practices; and practice greater transparency towards decision-makers.

Significance: By synthesizing the concepts and challenges surrounding RA, this paper seeks to raise awareness concerning the limitations of RA as a tool for decision-making. In addition, the recommendations are intended to encourage improved scientific practices and transparency when conducting RA.

Future Plans: This paper provides a brief review of only a select number of existing references on RA. In the future, it could be valuable to expand on this work in order to produce a more extensive review of the concepts, challenges, and recommendations for RA practices.

Sommaire

Évaluation des risques : Défis et recommandations

Kyungryun (Cathy) Pak, Lynne Genik, Centre des sciences pour la sécurité de RDDC; DRDC CSS TN 2012-032; R & D pour la défense Canada – CSS; Décembre 2012.

Contexte : Dans le cadre du projet de collaboration entre Recherche et développement pour la défense Canada (RDDC) et Emergency Management British Columbia (EMBC), des recherches documentaires ont été menées dans les domaines de l'évaluation des risques et des infrastructures essentielles. Les documents de recherche sont constitués d'environ 200 documents de référence pour chacun des domaines. Une fois la préparation de ces documents terminée, RDDC a décidé de s'appuyer sur son travail en effectuant une analyse documentaire relative à l'évaluation des risques dans le but de cerner les défis et les contraintes courants liés aux pratiques d'évaluation des risques et de formuler des recommandations. Ce document présente les résultats de l'analyse documentaire.

Méthodologie : L'analyse documentaire a été réalisée en s'appuyant sur les documents de référence compris dans la recherche documentaire de RDDC sur l'évaluation des risques. Dans une moindre mesure, des recherches supplémentaires ont également été effectuées, principalement au moyen de bases de données en ligne, notamment la base de données de RDDC et celle de la bibliothèque de l'Université d'Ottawa. L'auteur principal a choisi plusieurs documents de référence qui présentaient des analyses sur les concepts, les limites et les recommandations concernant l'évaluation des risques. Ces analyses ont permis de dégager et de présenter en synthèse les éléments clés pour donner un aperçu des facteurs importants liés à l'évaluation des risques.

Résultats : Ce document présente une synthèse des écrits existants sur l'évaluation des risques. Tout d'abord, on y décrit les méthodes quantitatives, qualitatives, intégrées et comparatives de l'évaluation des risques, en insistant sur les limites de ces méthodes. On aborde ensuite les répercussions des heuristiques psychologiques sur la capacité de l'être humain de porter des jugements exacts. Cela comprend l'heuristique de disponibilité, l'heuristique de représentativité et l'heuristique d'ancrage et d'ajustement. Enfin, on formule des recommandations à l'intention des spécialistes de la gestion du risque, c'est-à-dire cerner et communiquer l'incertitude, adopter des pratiques scientifique et favoriser la transparence à l'égard des décideurs.

Importance : Par une synthèse des concepts et des défis entourant l'évaluation des risques, ce document vise à mieux faire connaître les limites de l'évaluation des risques en tant qu'outil décisionnel. En outre, les recommandations ont pour but d'encourager de meilleures pratiques scientifiques et une meilleure transparence lors d'une évaluation des risques.

Perspectives : Ce document ne donne qu'un bref aperçu d'un certain nombre de documents de référence existants sur l'évaluation des risques. Dans le futur, il pourrait être utile de s'appuyer sur ces travaux afin de réaliser un examen plus approfondi des concepts, des défis et des recommandations relatifs aux pratiques d'évaluation des risques.

Table of Contents

1 Introduction

1.1 Background

The collaborative project between Defence Research and Development Canada (DRDC) and Emergency Management British Columbia (EMBC) was established with the aim of demonstrating the value of a scientific approach in improving EMBC's programs in the areas of risk assessment (RA) and critical infrastructure (CI). During the initial phases of this project, a high-level literature review was conducted for RA and CI as it relates to public safety and security [1]. Identifying the potential value of more extensive literature searches, the authors completed documented literature searches in these two areas [2, 3], each of which provide descriptions and a categorization scheme for approximately 200 references. With the completion of these literature searches, the authors decided to select several references that discuss the common issues and challenges of RA, and provide recommendations for improving RA practices. This paper synthesizes the results of this literature review, with the aim of raising awareness regarding the limitations of RA as well as the measures that must be taken in order to improve its value for decision-making.

1.2 Risk Assessment

Risk, as defined by the Society for Risk Analysis, is the "potential for realization of unwanted, adverse consequences to human life, health, property, or the environment" [4]. In order to manage risk, organizations and groups of all types and sizes coordinate activities that provide direction and control with regard to risk. A key element of risk management is RA, which comprises risk identification, risk analysis, and risk evaluation [5]. RA allows organizations to gain an enhanced understanding of risk, their causes, consequences and probabilities, and is used to "provide evidence-based information and analysis to make informed decisions on how to treat particular risks and how to select between options" [6]. Though RA emerged in its applications to engineering and health [7], its use is now widespread across a variety of fields, including that of public safety and security. As a result, RA can have great implications on the safety and security of affected populations, and it is critical to be educated in its use and application.

2 Risk Assessment Approaches

There is no unique method or approach for RA, as its diverse applications necessitate different methods that are most appropriate for the scope of the assessment and the availability of resources. This section will discuss the differences, benefits, and limitations of several RA approaches. Section 2.1 will provide a discussion on qualitative and quantitative RA methods, and Section 2.2 will explore integrated and comparative RA.

2.1 Qualitative and Quantitative Risk Assessment

One of the most widely debated issues in RA is the choice between using qualitative and quantitative methods. The sections below provide a brief overview of some of the benefits and limitations of these two approaches.

2.1.1 Qualitative Risk Assessment Methods

Qualitative RA methods assign verbal ratings such as “high”, “medium” or “low” to various dimensions of risk, such as probability[1] and consequence. They are frequently used because they simplify risk assessments, reduce the required inputs and judgements, and can be easily communicated to policy makers and stakeholders [8]. However, the simplicity of qualitative methods generates a variety of concerns with respect to mathematical accuracy and precision.

2.1.1.1 Frequency and Probability Ratings

For frequency and probability ratings, Dooley [9] explains that qualitative RA uses a few intervals to represent a wide and continuous range of frequencies. An example of frequency categories is illustrated below:

Frequency Category	Qualitative Definition
Frequent	Occurs continuously
Likely	Occurs at a high rate
Occasional	Occurs sporadically (irregularly, sparsely, or sometimes)
Seldom	Occurs rarely
Unlikely	Occurs very rarely

In the above example, each category represents a broad range of frequencies (that is, all of the events that “occur at a high rate” are represented by the category, “likely”). When using such intervals, the differences that exist within each category are not reflected, and precision is lost. In addition, these intervals leave great room for inaccuracy and error, for it is very difficult to distinguish between the categories in a consistent manner; the understanding and assignment of a category is very subjective and relies on individual judgement.

[1] In this paper, probability, frequency, and likelihood are used interchangeably.

2.1.1.2 Consequence Ratings

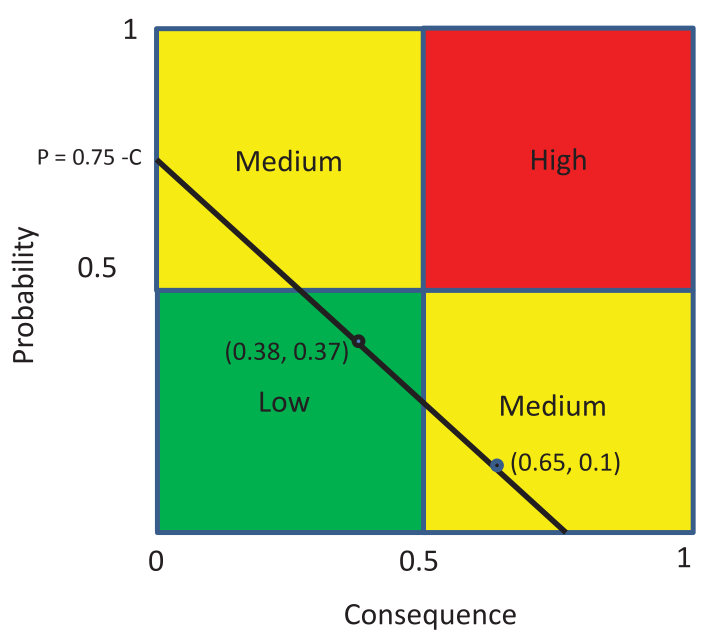
The problems discussed above apply equally to consequence ratings [9]. Furthermore, Cox, in *What's Wrong with Risk Matrices* [10], explains that there is no objective way to classify the relative severity of events that have highly uncertain consequences. Individuals have varying degrees of risk aversion (avoidance of risk), and these attitudes can change the way that humans order the consequence ratings of several events. A more detailed explanation of this hypothesis can be found in Cox's paper.

2.1.1.3 Risk Matrices

Cox also explores the challenges and errors that arise in the use of risk matrices. A risk matrix is a tool that is frequently used in order to assess and communicate risk. It is a table that assigns categories for frequency on its rows, and categories for consequence on its columns (or vice versa). Using the formula Risk = Probability x Consequence, risk matrices permit users to determine various levels of risks for each row-column pair.

The simplicity of risk matrices have led to their widespread use, and several national and international standards have recommended risk matrices to aid in setting priorities and guiding resource allocations. Many argue that risk matrices, though imprecise, provide value in assisting decision-makers to focus attention on the most serious problems. However, Cox demonstrates that risk matrices can sometimes provide information that is "worse than useless" [10]. This can occur when probability and consequence values are negatively correlated; that is, as one value increases, the other value decreases. This is illustrated in the following example.[2]

Suppose that a decision-maker wants to eliminate only one out of two risks.



[2] The values and points shown on this diagram are approximate.

In the risk matrix shown above, consequence and probability have quantitative values between 0 and 1 (inclusive). The quantitative risk for any pair (consequence, probability) is their product, as defined by the formula: Risk = Consequence x Probability.

In general, the risk matrix classifies the two risks with no error if one risk falls into the "high" risk cell, and the other falls into the "low" risk cell. This is because every risk in the high cell is both qualitatively and quantitatively greater than any risk that falls in the "low" risk cell.

However, suppose that probability and consequence values are negatively correlated (that is, as the consequence value increases, the probability value decreases), and that the risk pairs are concentrated along the line Probability = 0.75 – Consequence. For all of the risks that fall along this line, the risk classifications provided by the risk matrix do not accurately represent the true quantitative risks. Along this line, the risks that fall in the "medium" risk cells actually have *smaller* quantitative risks than the risks that fall in the "low" risk cells.

For example, consider two risks:
R_1 = ($Consequence_1$, $Probability_1$) = (0.38 and 0.37)
and R_2 = ($Consequence_2$, $Probability_2$) = (0.65, 0.1).

As shown on the matrix, the pair R_1 = (0.38, 0.37) falls in the "low" cell of the risk matrix, while the pair R_2 = (0.65, 0.1) falls in the "medium" cell of the risk matrix.

However, by applying the formula, Risk = Consequence x Probability, it is possible to calculate the quantitative risks of these two pairs.
R_1 = 0.38 x 0.37 = 0.14
R_2 = 0.65 x 0.1 = 0.065

In this case, the qualitatively "low" risk, R_1 actually has a (quantitatively) higher risk than the qualitatively "medium" risk R_2. Thus, the risk matrix reverses the correct (quantitative) risk ratings by assigning the higher qualitative risk category to the quantitatively smaller risk. Cox describes this situation as one in which the risk matrix is "worse than useless", for a decision-maker who uses the risk matrix would have been better off by making a decision randomly.

2.1.2 Quantitative Risk Assessment Methods

In order to overcome the limitations that emerge in qualitative RA, many authors suggest the use of quantitative methods. According to Kaplan and Garrick [11], quantitative RA is conducted by answering the following three questions:

1) What can happen?
2) What are the consequences?
3) What is the likelihood?

The first question is answered by identifying a set of initiating events or disturbances. Next, the consequences can be estimated by using logic diagrams such as event and fault trees to generate event sequences (scenarios). Lastly, the likelihood of these scenarios is evaluated using all available evidence (primarily historical data) and expert judgement and assigned numerical rankings according to the expected frequency of occurrence [12].

The use of quantitative RA is widespread because it employs logical and transparent mathematics and provides precise numerical values for risk [8]. In addition, quantitative RA is able to consider a large number of scenarios and quantify uncertainty [12]. However, the reliability and accuracy of quantitative RA can be challenged as a result of the numerous limitations and errors in quantitative RA methods.

2.1.2.1 Likelihood Ratings

Likelihood ratings can be a source of human error and uncertainty. In quantitative RA, likelihood is determined using precise quantitative interval scales, such as: very likely: >90%; likely: >66%; about as likely as not: 33% to 66%; unlikely: <33%; and very unlikely: <10%. However, such interval scales are often misinterpreted and erroneously applied in practice [13]. This is further explained in Section 2.1.2.3.

2.1.2.2 Consequence Ratings

The consequences of risk events are rated using a variety of "impact factors", such as fatalities, economic damage, environmental impact, etc. However, estimating the impacts of potential scenarios creates several challenges because some impact factors are very difficult to quantify and measure. For example, how does one quantify the damage to national reputation or environmental degradation? In addition, impact estimates must be aggregated into a common metric, so that an overall risk rating can be determined for each scenario. However, the "impact factors" are very diverse, and thus use different measurement units (such as the number of fatalities or the monetary loss in dollars). Consequently, the conversion of impact estimates into a common metric requires the use of subjective judgements. For example, risk assessors are often faced with the question, "what is the monetary value of a life"? [9].

2.1.2.3 Scoring Methods and Ordinal Scales

The consequence and likelihood ratings discussed above are frequently conducted using scoring methods with ordinal scales (described below). However, Hubbard and Evans, in *Problems with Scoring Methods and Ordinal Scales in Risk Assessment* [14], examine some of the limitations of this approach.

On an ordinal scale, items are assigned numbers in such a way that the order of the numbers represents the position or order of these items in relation to another. For example, on an ordinal likelihood scale of 0 to 5 (where 0 represents no likelihood and 5 represents very high likelihood), an item with a likelihood rating of 4 has a greater likelihood than an item given a rating of 2. Ordinal scales are used to rate risk factors, which are then aggregated using additive weighting or multiplication in order to provide a measure of the overall risk. The additive method is generally used to assess the overall risks of a project, policy, or investment. For example, if a manufacturer wants to evaluate the credit risk of an international customer, he or she might identify risk factors such as the country stability, order size, or currency volatility. Using the additive method, each of these risk factors would be weighted according to their importance, and then added to arrive at an overall risk score. The multiplicative method, on the other hand, is used to assess the risk of individual events. It is common to use scores for likelihood and consequence and multiply these values to produce a final risk rating. This multiplicative method is consistent with the commonly used definition of risk: risk = likelihood x consequence. After providing an overview of the existing research on this method, the authors claim that "scoring methods are not useful tools for risk assessment" [14].

Hubbard and Evans outline four major weaknesses of quantitative ordinal scales in RA. They are:

1. These methods are subject to a variety of cognitive biases which impair peoples' ability to accurately measure risk.

   Hubbard and Evans provide a brief literature review of psychological research studies concerning cognitive biases and other psychological phenomenon that skew human judgement. They are identified as follows:

- Availability Heuristic[3]: People tend to ignore statistical evidence and base their estimates on their memories, which favour the most vivid, emotionally charged, and unusual events.
- Gambler's Fallacy: People make the assumption that individual random events are influenced by previous random events.
- Optimism bias: People overestimate the probability that positive events will occur for them, in comparison with the probability that these events will occur for other people.
- Confirmation bias: People seek to confirm their preconceived beliefs when interpreting or gathering information.
- Self-serving bias: People claim more responsibility for successes than failures.
- Anchoring: People tend to base their estimates on previously conceptualized quantities, even when the two quantities are completely unrelated.
- Overconfidence: People consistently overestimate the certainty of their forecasts.
- Inconsistency: When asked to evaluate the same item on separate occasions, people tend to provide different estimates, despite the fact that the information has not changed.

A more detailed explanation of psychological heuristics and biases can be found in Section 3.

2. Verbal labels used in ordinal scales are interpreted with extreme inconsistency.

   Verbal labels for frequency and probability ratings are often used with the assumption that:
   - the simplicity of the labels will lead to improved comprehension and application; and
   - the quantitative descriptions of these labels will lead to the accurate and consistent application of the scales.

The authors refer to the studies of Budescu et al [13], which demonstrated that even when users were provided with detailed and quantitative definitions for likelihood scales, the users interpreted the labels differently when applying them to specific events or situations. These varied interpretations violated the provided definitions. In fact, it was even possible that the label, "unlikely" (defined as less than 33%), was interpreted to mean "as much as a 66% probability".

3. The features of the scoring scheme have a large impact on the results of RAs.

   Firstly, Hubbard and Evans claim that RA users tend to treat ordinal scales as if they are ratio scales. Unlike in ordinal scales, ratio scales assign numerical values to items in such a way that the differences between these values are significant. For example, on a ratio scale for likelihood, a rating of 4 represents a likelihood that is twice as great as the likelihood of an item assigned a rating of 2. Hence, if an ordinal scale gives "not very likely" a numerical value of 2 and "very likely" a value of 4, someone who is using this ordinal scale as a ratio scale may wrongly assume that "very likely" is two times more likely than "not very likely". These invalid assumptions can have adverse impacts on the results of the RA. In addition, ordinal scales tend to only use a small number of discrete values (for example, a range from 0 to 5). As a result, each scale represents an extremely wide range of values, which can lead to the loss of resolution. Lastly, the authors highlight the inadequacy of simple additive or multiplicative methods to accurately represent risks.

4. Ordinal scoring methods rarely consider the correlations between risks, which can lead to the severe underestimation of risks.

   According to Hubbard and Evans, the multiplicative model (risk = likelihood x consequence) assumes that the likelihoods and consequences of different risk events are entirely independent.

[3] A heuristic is a strategy, trick, simplification, or any other device that allows people to more efficiently find solutions in large and complex problem spaces [15].

This model may, for example, determine that the risks of three different events are "low" or "medium". However, this method does not account for the correlations that may exist between different events, or the possibility that a single event can cause a chain reaction of other events (known as "cascade failures"). These "low" or "medium" risks may have correlations and interdependencies that, when considered together, may produce one very high risk. Therefore, the Risk = Likelihood x Consequence model can severely underestimate risks by falsely assuming the independence of different risk events.

In view of the problems and limitations discussed above, Hubbard and Evans suggest that RA should meet the following three criteria:

1. RA should use quantitative expressions of probability and the magnitude of losses. For example, instead of stating that likelihood or impact are "medium" or "high", one should state, "10% chance of a loss of inventory worth $2 million". These methods are very complex and resource-intensive, however, and will not be discussed within this paper.
2. RA should explicitly model the relationships and correlations between different components of a system. Instead of adding or multiplying risk factors, risk should be represented using descriptive functions.
3. RA should apply methods to correct the biases in human judgement and decision-making.

2.2 Integrated and Comparative Approaches to Risk Assessment

The diverse forms and causes of risk create challenges in conducting risk assessments. Risks can take the form of non-malicious hazards (such as natural disasters and health hazards), and malicious or human-caused threats. Thus, in order to effectively manage risks, it is necessary to consider all hazards and threats. However, all-hazards RA can be conducted using a variety of different approaches. This section provides a brief discussion on the differences, benefits and limitations of integrated and comparative approaches to all-hazards RA.

2.2.1 Integrated Risk Assessment

A commonly used method for all-hazards RA is the integrated approach. This approach requires the collection of risk analyses for all possible hazards and threats, and aggregates them into one holistic and complete analysis using a common metric.

All-hazards RA is advantageous because it allows all considered risks to be ordered in relation to one another. As a result, the most significant risks can be identified in order to support more effective decision-making. However, in its *Review of the Department of Homeland Security's Approach to Risk Analysis* [16], the United States (US) National Research Council (NRC) highlights the challenges of using an integrated approach to RA. Since the risks posed by various hazards or threats are very different in nature, the methods employed to assess these risks are varied and diverse. For example, RA for natural hazards is supported by a wealth of historical data that allows analysts to model the risk events and make rather accurate estimations on risk factors, such as the location or frequency of occurrence. On the other hand, terrorist attacks are carried out by intelligent actors and may have no historical precedent. As a result, the risk of terrorist attacks is very difficult to predict. The differences in the assumptions and methods used for these diverse risk events make it very challenging and impractical to integrate them into one common metric. Therefore, the NRC proposes the use of comparative risk analysis.

2.2.2 Comparative Risk Assessment

Similarly to integrated RA, comparative RA performs risk analyses for all of the potential risks that are faced by an entity or community. However, instead of aggregating these risks into a common metric, comparative RA uses qualitative methods to compare and contrast these risks. The NRC suggests that methods such as Delphi analysis and expert judgement can provide valuable comparisons which can support decision making. Although this approach compares across a variety of different metrics, the NRC claims that "the scope and diversity in the metrics can themselves be very informative for decision making" [16].

3 Psychological Considerations

Despite the rigorous mathematical formulations and methodologies that have been developed in order to improve the objectivity and accuracy of RA, it is important to recognize that RA is conducted by (imperfect) humans. As explained by Rosqvist [17], the inputs of domain experts are required during the hazard/threat identification, risk estimation, and risk evaluation phases of RA. However, research findings demonstrate that human judgements are influenced by a variety of psychological biases and heuristics which compromise peoples' ability to provide accurate or reliable estimations. Thus, this section will provide a brief discussion of three prominent psychological heuristics by referring to the foundational work of Tversky and Kahneman.

3.1 Availability Heuristic

The availability heuristic is used when people assess the frequency or probability of an event by drawing from incidents that are most easily brought to mind. In a study conducted by Tversky and Kahneman [18], participants were asked to compare the likelihoods of two occurrences.

> Scenario A: In typical English text, a word (of three or more letters) will start with the letter k.
> Scenario B: In typical English text, a word (of three or more letters) will have k as its third letter.

More than 2/3 of participants estimated that scenario A is more likely than scenario B, despite the fact that scenario B is actually two times more likely than Scenario A. Tversky and Kahneman explain that people chose option A because it is easier to think of words that begin with k than those that have k as its third letter. Thus, the authors conclude that the respondents based their likelihood judgements based on what is easier to bring to mind.

3.2 Representativeness Heuristic:

The representativeness heuristic appears when people are faced with probabilistic questions such as:

1. What is the probability that object A belongs to class B?
2. What is the probability that event C originates from process D?
3. What is the probability that process E will generate event F?

In answering these types of questions, people tend to use the representativeness heuristic, in which they assess the probability of occurrence based on the similarity of the two things (in the examples above, of A to B, C to D, and E to F).

For example, in a study conducted by Tversky and Kahneman [19], participants were given a description of a woman (Linda) that suggested that she was a feminist. The participants were then asked to assess the likelihood of the following two scenarios:

> Scenario A: Linda is a bank teller.
> Scenario B: Linda is a bank teller and a feminist.

In the results, 80-90% of respondents considered scenario B to be more likely than scenario A. However, the co-occurrence of two events cannot be more likely than the occurrence of either event alone. Nevertheless, most participants chose the second scenario because it is more similar to the description provided. Hence, the representativeness heuristic causes people to confuse similarity with probability.

3.3 Anchor-and-Adjustment Heuristic

Anchoring and adjustment is used when a quantitative estimate is made by starting from an initial value and adjusting it up or down in order to arrive at the final answer. This initial value may be given in the formulation of the problem, or may be the result of a partial computation. However, people typically underestimate the adjustments that are needed to arrive at the correct value.

For example, in a study conducted by Tversky and Kahneman [20], two groups of people were asked whether or not the percentage of African countries in the United Nations was greater or less than a given value. This value was chosen randomly using a spinning wheel of fortune that was turned in the respondents' presence. The first group was given 10% as their initial value and the other group was given 65%. These groups were then asked to estimate the actual percentage of African countries in the United Nations. The first group had a median estimate of 25% and the second group had a median estimate of 45%. Hence, the numbers arbitrarily chosen by the spinning wheel had a significant effect on the respondents' answers, as they "anchored" their responses to the initial values and "adjusted" them to arrive at their final estimate.

In view of the biases and uncertainties that are embedded in RA processes, risk practitioners should exercise greater caution in the use and application of RA. The subsequent section provides recommendations for measures that can be taken in order to improve the value and reliability of RA.

4 Moving Forward: Recommendations

This paper has presented some of the problems and challenges that arise when conducting RA. In order to ensure that RA is an effective tool for decision-making, risk practitioners should be aware of these issues and adopt appropriate measures. In its evaluation of the Department of Homeland Security's Approach to Risk Analysis, the US NRC provided several recommendations for improving the scientific rigour of RA [16]. These are described below.

4.1 Uncertainties

RA practitioners attempt to understand the events of the past and predict the future. By the very nature of the task, RA is subject to a number of uncertainties. Uncertainties arise from incomplete information or data; limited understanding of environmental and human responses; subjectivity in threat, vulnerability, or consequence analyses; and the difficulties in integrating existing knowledge into models that can provide predictions.

Uncertainty arises in two different forms:

- *Aleatory or statistical uncertainty:* inherent uncertainties that arise from the random variations that occur across time, space, or from sample to sample. An example is the variability of the results when flipping a coin.
- *Epistemic or systematic uncertainty*: uncertainty that is the result of a lack of knowledge about the system. This type of uncertainty can be reduced.

Both aleatory and epistemic uncertainty must be recognized, characterized, and communicated throughout all of the steps of RA. As outlined in the ISO 31000, effective risk management "explicitly takes account of uncertainty, the nature of that uncertainty, and how it can be addressed" [5]. In view of these uncertainties, risk practitioners should also be careful to avoid false precision. Placing too much emphasis and trust on quantifying RA can be misleading if the levels of precision and uncertainty surrounding the numbers are not properly communicated.

4.2 Scientific Practices

In view of these uncertainties, the NRC recommends that risk analysts adopt critical scientific practices such as documentation, validation, peer review, and publishing. These measures allow others to determine the reliability and usefulness of RA as an input to decision-making. These recommendations are consistent with the principles of the ISO 31000, which state that effective risk management "is based on the best available information" [5].

The NRC emphasizes that it is important to evaluate the quality of RA by assessing the logical purpose and structure of the risk model, as well as the inputs and assumptions that affect its application. The NRC provides the following requirements for good scientific practice for model-based scientific work:

- "Clear definition of model purpose and decisions to be supported;
- Comparison of the model with known theory and/or simple test cases or extreme situations;
- Documentation and peer review of the mathematical model, generally through a published paper that describes in some detail the structure and mathematical validity of the model's calculations; and

- Some verification and validation steps, to ensure that the software implementation is an accurate representation of the model and that the resulting software is a reliable representation of the model and that the resulting software is a reliable representation of reality" [16].

Furthermore, the NRC refers to its 2008 report, *Department of Homeland Security Bioterrorism Risk Assessment: A Call for Change* [21] in order to make recommendations for effective documentation. The following elements should be included in the documentation of RA processes:

1. "How they construct risk assessment models;
2. What assumptions are made to characterize relationships among variables and parameters and the justification for these;
3. The mathematical foundations of the analysis;
4. The sources of the values assigned to the parameters for which there are no available data; and
5. The anticipated impact of uncertainty for assumptions and parameters" [21].

4.3 Improving Transparency

Recognizing that RA is a tool for decision-making, risk analysts must ensure that the results of RA are credible and informative for decision makers. As a result, RA must be transparent, and information concerning the assumptions, the quality of the data inputs, and the anticipated uncertainties must be communicated. This allows decision makers to make more informed decisions when comparing a broad range of risks, especially for those involving high levels of uncertainty. These recommendations are further supported by the ISO 31000 principles, which stipulate that effective risk management is "transparent and inclusive" [5]. The ISO 31000 adds that stakeholders and decision makers should be properly represented and involved when determining the risk criteria.

5 Summary

This paper presents the results of a literature review of risk assessment (RA), which was undertaken as a subsequent step to the literature search on RA that was conducted for the DRDC-EMBC collaborative project. In order to build upon this literature search, the authors identified a few key references and synthesized the common issues, challenges, and recommendations for RA. The characteristics and limitations of qualitative, quantitative, integrated, and comparative RA methods are presented. The paper also describes the availability, representativeness, and anchor-and-adjustment heuristics, and discusses their implications on human judgement. In view of these challenges and limitations, risk practitioners are advised to identify and communicate uncertainty, adopt scientific practices, and increase their transparency towards decision-makers.

6 References

[1] Genik, L., & Chouinard, P., *DRDC Support to Emergency Management British Columbia's (EMBC) Hazard Risk Vulnerability Analysis (HRVA) and Critical Infrastructure (CI) Programs: Problem Formulation and Solution Strategy*, DRDC Technical Memorandum, October 2012, DRDC CSS TM 2012-015

[2] Pak, K., & Genik, L., *Risk Assessment References: Documented Literature Search*, DRDC Technical Note, October 2012, DRDC CSS TN 2012-014.

[3] Pak, K., & Genik, L., *Critical Infrastructure References: Documented Literature Search*, DRDC Technical Note, October 2012, DRDC CSS TN 2012-013.

[4] *Risk Analysis Glossary*, Society for Risk Analysis, available from http://www.sra.org/resources_glossary_p-r.php

[5] *Risk Management – Principles and Guidelines on Implementation*. International Organization for Standardization. ISO 31000, 2009, First Edition.

[6] *Risk Management – Risk Assessment Techniques.* International Electrotechnical Commission. IEC/FDIS 31010, 2009, Final Draft.

[7] Paté-Cornell, M.E. (1996). *Uncertainties in Risk Analysis: Six Levels of Treatment*, Reliability Engineering and System Safety, 54: 95-111.

[8] Cox, L.A., Babayev, D., & Huber, W. (2005). *Some Limitations of Qualitative Risk Rating Systems*. Risk Analysis, 25(3): 651-662.

[9] Dooley, P., *Risk Analysis: Concepts, Practices, and Challenges*, Slideshow, 19 January 2012.

[10] Cox, L.A. (2008). *What's Wrong with Risk Matrices?* Risk Analysis, 28(2): 497-512.

[11] Kaplan, S., & Garrick, B.J. (1981). *On the Quantitative Definition of Risk*. Risk Analysis, 1(1): 11-27.

[12] Apostolakis, G.E. (2004). *How Useful Is Quantitative Risk Assessment?* Risk Analysis, 24(3): 515-520.

[13] Budescu, D.V., Broomell, S., & Por, H. (2009). *Improving Communication of Uncertainty in the IPCC Reports*. Psychological Science,20 (3): 299-308.

[14] Hubbard, D., & Evans, D. (2010). *Problems with Scoring Methods and Ordinal Scales in Risk Assessment*. IBM Journal of Research and Development, 54(3): 1-10.

[15] Feigenbaum, E.A., & Feldman, J. (Editors). (1963). *Computers and Thought*. McGraw-Hill Inc., New York.

[16] Committee to Review the Department of Homeland Security's Approach to Risk Analysis: National Research Council. *Review of the Department of Homeland Security's Approach to Risk Analysis*, The National Academies Press, 2010, 160 pages. Available from www.nap.edu/catalog.php?record_id=12972.

[17] Rosqvist, T. (2003) *On the Use of Expert Judgement in the Qualification of Risk Assessment*. VTT Publications 507, Espoo, Finland.

[18] Tversky, A., & Kahneman, D. (1973). *Availability: A Heuristic for Judging Frequency and Probability*. Cognitive Psychology, 5: 207 – 232.

[19] Tversky, A., & Kahneman, D. (1982). Judgements of and by Representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and Biases* (pp. 84 – 98). Cambridge, UK: Cambridge University Press.

[20] Tversky, A., & Kahneman, D. (1974). *Judgement under Uncertainty: Heuristics and Biases.* Science, 185, 1124 – 1131.

[21] Committee on Methodological Improvements to the Department of Homeland Security's Biological Agent Risk Analysis: National Research Council. *Department of Homeland Security Bioterrorism Risk Assessment: A Call for Change,* National Academies Press, 2008, 173 pages.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, for example Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Centre for Security Science Defence R&D Canada 222 Nepean St. 11th Floor Ottawa, ON Canada K1A 0K2	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED NON-CONTROLLED GOODS DMC A Review: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Risk Assessment: Challenges and Recommendations

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

Pak, K

Genik, L

5. DATE OF PUBLICATION (Month and year of publication of document.) December 2012	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 21	6b. NO. OF REFS (Total cited in document.) 21

7. DESCRIPTIVE NOTES (The category of the document, for example technical report, technical note or memorandum. If appropriate, enter the type of report, for example interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Note

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Centre for Security Science

Defence R&D Canada

222 Nepean St. 11th Floor

Ottawa, ON Canada K1A 0K2

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) File 3700-1	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC CSS TN 2012-032	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.)) Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

This paper presents a literature review on risk assessment (RA) that focuses on some of the challenges that arise when conducting RAs. This work builds upon a previous literature search on RA, identifying several key references and synthesizing common issues and recommendations. An assessment of quantitative, qualitative, integrated and comparative RA approaches and a discussion on availability, representativeness and anchor-and-adjustment psychological heuristics are included. The following recommendations are provided: that risk practitioners recognize and communicate uncertainty; adopt scientific practices such as documentation, validation, peer review, and publishing; and improve transparency. The literature review was undertaken as a part of the Defence Research and Development Canada (DRDC) and Emergency Management British Columbia (EMBC) collaborative project and is intended to be a resource for DRDC, EMBC, and other partners.

Le présent document contient une analyse documentaire sur l’évaluation des risques, axée sur certains des défis qui se présentent lors des évaluations des risques. Ce travail s’appuie sur une recherche documentaire antérieure sur l’évaluation des risques qui a permis de repérer plusieurs documents de référence clés et de faire une synthèse des questions et des recommandations courantes. Ce document comprend une évaluation des méthodes quantitatives, qualitatives, intégrées et comparatives d’évaluation des risques ainsi qu’une discussion sur les heuristiques psychologiques (heuristique de disponibilité, de représentativité et d’ancrage et d’ajustement). Les recommandations suivantes ont été formulées : que les spécialistes de la gestion du risque reconnaissent et communiquent l’incertitude, qu’ils adoptent des pratiques scientifiques comme la documentation, la validation, l’évaluation par les pairs et la publication et qu’ils améliorent la transparence. L’analyse documentaire a été entreprise dans le cadre d’un projet de collaboration entre Recherche et développement pour la défense Canada (RDDC) et Emergency Management British Columbia (EMBC). Il s’agit d’un outil que RDDC, EMBC et d’autres partenaires peuvent utiliser comme ressource.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, for example Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Risk Assessment; Literature Review; Challenges, Recommendations, Quantitative Risk Assessment, Qualitative Risk Assessment, Integrated Risk Assessment, Comparative Risk Assessment, Availability Heuristic, Representativeness Heuristic, Anchor-and-Adjustment Heuristic, Uncertainty, Scientific Practices, Transparency