HAL open science

The Illusion of Risk Control

Gilles Motet, Corinne Bieder

▶ To cite this version:

Gilles Motet, Corinne Bieder (Dir.). The Illusion of Risk Control. Springer, 2017, 978-3-319-32938-3. hal-01576144

HAL Id: hal-01576144

https://enac.hal.science/hal-01576144v1

Submitted on 21 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

SPRINGER BRIEFS IN
APPLIED SCIENCES AND TECHNOLOGY · SAFETY MANAGEMENT

Gilles Motet
Corinne Bieder *Editors*

The Illusion of Risk Control

What Does it Take to Live With Uncertainty?

FONCSI
Fondation pour une culture de sécurité industrielle

Springer Open

SpringerBriefs in Applied Sciences and Technology

Safety Management

Series Editors

Eric Marsden, Toulouse, France
Caroline Kamaté, Toulouse, France
François Daniellou, Toulouse, France

The SpringerBriefs in Safety Management present cutting-edge research results on the management of technological risks and decision-making in high-stakes settings.

Decision-making in high-hazard environments is often affected by uncertainty and ambiguity; it is characterized by tradeoffs between multiple, competing objectives. Managers and regulators need conceptual tools to help them develop risk management strategies, establish appropriate compromises and justify their decisions in such ambiguous settings. This series weaves together insights from multiple scientific disciplines that shed light on these problems, including organization studies, psychology, sociology, economics, law and engineering. It explores novel topics related to safety management, anticipating operational challenges in high-hazard industries and the societal concerns associated with these activities.

These publications are by and for academics and practitioners (industry, regulators) in safety management and risk research. Relevant industry sectors include nuclear, offshore oil and gas, chemicals processing, aviation, railways, construction and healthcare. Some emphasis is placed on explaining concepts to a non-specialized audience and the shorter format ensures a concentrated approach to the topics treated.

More information about this series at http://www.springer.com/series/15119

Gilles Motet · Corinne Bieder
Editors

The Illusion of Risk Control

What Does it Take to Live With Uncertainty?

Springer Open

Editors
Gilles Motet
INSA. LAAS-CNRS
Université de Toulouse
Toulouse
France

Corinne Bieder
Ecole Nationale de l'Aviation Civile (ENAC)
Toulouse
France

ISSN 2191-530X ISSN 2191-5318 (electronic)
SpringerBriefs in Applied Sciences and Technology
ISSN 2520-8004 ISSN 2520-8012 (electronic)
Safety Management
ISBN 978-3-319-32938-3 ISBN 978-3-319-32939-0 (eBook)
DOI 10.1007/978-3-319-32939-0

Library of Congress Control Number: 2016939387

© The Editor(s) (if applicable) and The Author(s) 2017. This book is an open access publication
Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.
The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

1 Uncertainty: New Perspectives, Questions and Proposals 1
Gilles Motet

2 Uncertainty—Its Ontological Status and Relation to Safety 5
Ove Njå, Øivind Solberg and Geir Sverre Braut

3 A Conceptual Foundation for Assessing and Managing Risk, Surprises and Black Swans 23
Terje Aven

4 Recognizing Complexity in Risk Management: The Challenge of the Improbable 41
Jean Pariès

5 Practices in the Danger Culture of Late Industrial Society. 57
Arie Rip

6 Judicial Review of Uncertain Risks in Scientific Research 67
Eric E. Johnson

7 What Can Japan's Early Modern Capital of Edo Teach Us About Risk Management?. 87
Jordan Sand

8 Conclusion 107
Corinne Bieder

Chapter 1
Uncertainty: New Perspectives, Questions and Proposals

Gilles Motet

Abstract Safety is ensured when hazardous circumstances, their consequences and their controls are known. How then, in uncertain contexts can safety be defined and achieved? Uncertainty raises new questions and needs new approaches for dealing with safety. These are introduced in the first chapter and developed throughout the book.

Keywords Uncertainty · Safety

1.1 Uncertainty: A New Perspective on Safety

The term "technological risk" was initially associated with the occurrence of unwanted circumstances (e.g. leading to accidents) which are both possible and probable. "Possible" expresses the fact that these circumstances are known and enumerable; "probable" means that these circumstances may arise in the future. These circumstances are described, for instance, by combinations of events or by sequences of events (scenarios), or both. In the context of safety, the considered circumstances inevitably lead to harmful events (accidents) when these circumstances arise (the deterministic approach to risk assessment). If all the circumstances are known, treatments can be defined a priori. Barriers prevent the occurrence of accidents by avoiding the occurrence of undesirable combinations or by neutralizing dreaded sequences of events. In this paradigm, safety is defined by the absence of such circumstances and by the absence of accidents, and is guaranteed by the presence of risk controls (barriers).

An adaptation of this viewpoint, called the "probabilistic approach to risk assessment", involved a change to the assumption that the circumstances considered by risk analysts always lead to accidents, by attaching a probability to the appearance of harm. Furthermore, the harm caused is no longer assumed to be constant for the same circumstances. For example, the failure of a system component does not always lead to the system's failure; the consequences of the failure may be more or less serious. Risk, or rather its estimation, is then defined using criteria such as a combination of the

G. Motet (✉)
INSA. LAAS-CNRS, Université de Toulouse, Toulouse, France
e-mail: gilles.motet@insa-toulouse.fr

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_1

probability of the occurrence of a harmful event and the severity of the harms. Safety is defined by the absence of unacceptable risk, usually expressed as a threshold (e.g. probability of a catastrophic failure $<10^{-9}$ per flight hour). This viewpoint is the most commonly used today. Safety is still achieved by the introduction of barriers, whose availability (measured by the probability of failure) intervenes in the evaluation of residual risk. This value is compared with the maximum level of risk allowed (the acceptability threshold). While all accidents are not avoided, their risks are controlled.

However, these two viewpoints on risk and safety are only variations on the same underlying paradigm: the circumstances that may lead to accidents are known, as are their possible effects: causes and consequences are enumerable, even if they are probable and not certain. The risk controls (and therefore the level of safety) can be defined a priori. If unforeseen circumstances arise, they are handled by "experience feedback" and added to the list of circumstances, which is considered as a finite set.

We are currently facing a challenge to this founding paradigm of risk and safety, having to admit that the circumstances (such as initiating events and scenarios) which may lead to accidents are uncertain and potentially infinite. Our ignorance of causes (circumstances), effects (harms) and their relationships is primarily quantitative in nature. This ignorance concerns, for instance, the difficulty in determining the probability of occurrence of events (causes and effects), the effectiveness of the barriers and the severity of consequences. Our ignorance is also, even more importantly, qualitative. This concerns, for example, our inability to establish an exhaustive list of circumstances which may lead to accidents, as this list is not finite in complex systems. Similarly, the nature of the effects of certain circumstances (in particular medium- and long-term effects) cannot be predicted given the state of knowledge (e.g. emerging risks related to innovation). This paradigm shift has impacts on the concept of safety which has to be revisited, but also other related concepts such as risk acceptability and their operational implementations (models, techniques, processes, practices, cultures, etc.).

Until now, most approaches have aimed at reducing and finally liquidating uncertainty, being based on the illusory hope of the contribution of knowledge development. Other approaches consist in forbidding any circumstances that lead to increased uncertainty (e.g. certain applications of the precautionary principle). They have all proven unsuccessful, leading to dramatic accidents. The inability to nullify this uncertainty and the need to live with it must be accepted: such is the new premise.

However, the results of uncertainties on safety must not be considered as inevitable effects of fate that we have to suffer. Uncertainty gives a new perspective on safety.

1.2 Uncertainty: New Questions for Safety Management

Risk management has long been considered the magic wand to address safety. The growth of knowledge would lead to the disappearance of uncertainties and the development of fully controlled risks. This would promise the achievement of our objective: a safe world. Yet, it seems that in a number of accidents, risk management

was not up to the challenge. What challenge? Dealing with the unexpected? Yet, the unexpected is now acknowledged as being unescapable...

Before going any further, it is worth going back to the notion of unexpected. Why is the unexpected? Several philosophies exist. Because no one put sufficient effort into anticipation is one of them. No one had the proper knowledge to anticipate some scenarios is another one. Who by the way is a question that would be worth addressing? Is it the first line operator who faced the real-time situation? Or the organization that provides the operators with the means and conditions to do their job? Or the risk manager who is in charge of identifying all risks and reducing the unacceptable ones? Or everyone, each one at his/her level and within his/her scope?

Going back to the initial question of the origin of the unexpected, is it a crisis of means allocation or of profound beliefs...? Or illusions? The illusion that everything can be controlled, including all risks.

To describe or explain these new situations that escape traditional risk management approaches, a new "magic" concept was introduced: uncertainty.

Beyond the magic word, what does the concept mean? What can it describe... and not describe?

Has dealing with uncertainty ever been within the scope of traditional risk management methods? If yes, what type of uncertainty did it pretend to address?

Are there alternative approaches that would work beyond the scope of traditional risk management?

If yes, how do these approaches relate to risk analysis?

Should we oppose them? Can they coexist? Under what conditions? Can we envisage their complementarity? How to build it?

1.3 Uncertainty: New Proposals

This book addresses various facets of the previous questions, proposing several viewpoints considered by various disciplines.

In Chap. 2, Ove Nja discusses the concept of uncertainty, analyzing its ontological status and its connections to safety. The chapter introduces the whole rationality spectrum (from relativism to positivism) of the uncertainty concept, focusing on its impact on the safety concept. This introduction also discusses a number of philosophical issues. The purpose is illustrated on safety concerns of the health sector and civil aviation.

The three following chapters deal with the control of effects of uncertainty.

In Chap. 3, Terje Aven presents and discusses recent advances in the risk field faced with uncertainty, linked to the conceptualization of the risk and specially addressing unforeseen events, surprises and so-called black swans. It shows how the traditional probabilistic approaches can be extended for assessing and then for handling uncertainties in the safety domain.

In Chap. 4, Jean Pariès extends the previous chapter discussing the way safety can be managed in an uncertain context. He highlights the illusion of reducing uncertainty

and introduces and compares two management alternatives: Resilience engineering and High Reliability Organizations.

Nowadays, society through regulation requires safety cases. In a well-defined context, evidence must be provided to obtain an authorization to operate hazardous activities. Uncertainty disrupts this principle. In Chap. 5, Arie Rip develops, in a sociological perspective, the idea that regulation and other a priori risk controls are all part of the "danger culture" of our industrial society which requires conformance to rules for avoiding accidents. The chapter shows that uncertainty simultaneously jeopardizes this approach and, by allowing "gray zones" in which the actors can manoeuvre, provides some necessary flexibility to the system.

Chapter 6 examines risk governance in an uncertain context from a legal perspective. Eric E. Johnson considers the catastrophic possibility that the superconducting synchrotron particle accelerator built at CERN could create a black hole which would swallow the Earth. Courts have the power to order the halt of such an activity, but their usual approach to decisions on risky activities, based on expert judgment and benefit-cost criteria, are difficult to apply in situations in which possible consequences are properly cataclysmic and experts are all biased. The chapter analyzes a number of court decisions to extract general opinions about the way justice considers uncertainty in the safety domain.

Chapter 7 gives a historian's perspective. Jordan Sand compares the way in which the Japanese handled fires during the Edo period (from 1600 to 1868) and their present approach to technological risks, with a particular focus on the Fukushima nuclear disaster. The comparison reveals that the emphasis during the Edo period on strength and continuity of the social order rather than on the preservation of material property produced a different view of risk and uncertainty.

A last chapter concludes the book, summarizing the lessons learned and highlighting some open questions.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 2
Uncertainty—Its Ontological Status and Relation to Safety

Ove Njå, Øivind Solberg and Geir Sverre Braut

Abstract The concept of uncertainty is difficult to comprehend, even when we restrict our focus to safety science. In a world with various scientific philosophical stances, "uncertainty" is debated in various contexts. However, in an effort to go deeper into a more basic understanding of uncertainty our knowledge is quickly challenged. What exists? How do we know what exists? What can we know about it? Aiming these questions at uncertainty reveals that interpreting uncertainty as existing in any ontological sense is difficult to defend. Does this imply that uncertainty can only be understood in an epistemological sense or merely as a construct? Epistemological understandings of uncertainty encompass, in principle, the whole rationality spectrum from relativism to positivism, thus not excluding any form of analyses or understanding of uncertainty. However, we recognize the need for an increased understanding of which elements the uncertainty concept comprises, and possible consequences of an unreflective discarding of elements. Within the framework of a linear time concept consisting of the past, the present and not least the future, we claim that uncertainty's ontological status exists on various levels. In the present uncertainty is a purely epistemological category, and in the past uncertainty has its meaning related to what has been observed, recognized and comprehended, thus a methodological challenge. In the futuristic perspective uncertainty exists and cannot be reduced.

Keywords Uncertainty · Ontology · Epistemology · Safety · Risk

The copyright to third-party material quoted in this chapter is property of its respective owners and is not included in the Creative Commons license for this chapter.

O. Njå (✉) · Ø. Solberg · G.S. Braut
University of Stavanger, Stavanger, Norway
e-mail: ove.njaa@uis.no

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_2

2.1 Introduction

The rational actor paradigm is often used as a basis for decision making under uncertainty. Proponents of this paradigm consider the uncertainty concept to be something outside the risk concept [10]. Other researchers include the concept of uncertainty as an integral part of the risk concept, where risk is defined as "observable quantities and their associated uncertainties" under the context of a system or activity studied [2, 3]. Opposed to this there are researchers who relate uncertainty to accuracy of e.g. models, data and expert judgments [13]. However, there are very few concrete definitions of uncertainty applied in risk and safety analyses. Lipshitz and Strauss [11] offer interesting thoughts within the Naturalistic Decision Making-Paradigm, in which they relate uncertainty to individuals' cognitive reasoning based on: inadequate understanding, incomplete information or undifferentiated alternatives. They claim that these uncertainties lead to different coping strategies amongst people involved in crisis management. Furthermore Taleb's "black swan" metaphor has raised a vast number of studies trying to develop methodologies to cope with the phenomena of unknown unknowns.

A question worth raising is therefore: Does the safety, risk and uncertainty concepts applied to various infrequent phenomena make any sense in scientific perspectives?

2.1.1 From Risk to Uncertainty

Althaus [1] has presented a thorough analysis of various scientific disciplines' perspectives on risk as depicted in Table 2.1.

Table 2.1 Extract of the analysis presented by Althaus

Discipline	View on risk and safety	Knowledge applied to the unknown (domain)
Logic and mathematics	A calculable phenomenon	Calculations and modeling (method)
Science and medicine	An objective reality	Principles, postulates and calculations (ontology, epistemology and method)
Social sciences	A cultural or societal phenomenon	Culture, social constructs, (epistemology and method)
Economics	A decisional phenomenon	Decision-making theory and principles (epistemology and methods)
History and the humanities	As a story	Narrative (epistemology)
Philosophy	A problematic phenomenon	Wisdom (ontology, epistemology)

This table use a fairly large brush on each of the disciplines and treat them as uniform. Within the single disciplines there are of course different schools of thought and discussions, contributing to the continued development of each discipline.

For each discipline Althaus describes the related view on risk and what types of knowledge is applied to cope with what is termed the unknown. The brackets in the right column are added here as guide to which domain of world (or philosophy and science) this knowledge is related to according to our view. The different domains used are ontology, epistemology and method. Ontology and epistemology as presented in this article, and method as a domain not necessarily adhered to either ontology or epistemology, but an instrumental phenomenon aiming to produce valid argumentations.

When we use the term risk we relate to possible future events, something that has not happened yet, but might happen in the future. This something that might or might not happen often refers to an explicit danger and equally often does refer to an implicit reward. Typical examples of this can be: *the risk of driving, the risk of flying, the risk of oil exploration in Northern Norway, the risk of transporting dangerous goods* and so on. The challenge with the concept of risk is that risk is about the future. That is probably also the main reason for all the controversies and disputes over the concept of risk. Mastery of risk is the mastery of the all-time riddle: what will the future be like? In modern times, this has been modified to: what is most probable that the future would look like?

The main problem with this riddle is that our knowledge about the future in principle is zero. Anything could happen in the next instance. In current everyday language, many think of risk as synonymous with danger in the future, something to be avoided. This danger can also be seen as positive e.g. as a necessary component in an adventure. Another connotation that can be seen frequently is the adverbial risky with a meaning of probable; as in "it is risky", meaning it is very likely to occur. For others risk can be a business opportunity that can be exploited to maximize gains. We also speak about different types of risk e.g. objective, subjective and perceived risk.

Risk has become an increasingly important concept in the globalized world. From being based on individual and locally based assessments, risk is now a central part of global discussions and a driver of research on life threatening issues, such as climate change, terrorism and epidemics. The concept of risk is used as a tool for experts claiming to possess superior knowledge of future threats.

However, risk deals with the core of probability theory. The concept of probability is clarified by adopting Pierre Laplace's formulation of the mathematical theories of probability; "probability theory is only good sense reduced to calculus". When theories of probability conflicted with the intuition of "reasonable men", mathematicians went back to the drawing board and changed their theories: a clash between probability theory and intuition meant the theory was wrong, not the intuition.

Uncertainty in relation to the findings from risk analyses has become a much debated topic in the field of risk research. Pending how uncertainty is understood and used in risk analyses and risk management, uncertainty can be seen as something we can deal with and consequently reduce. The presumption is that if we have more knowledge, uncertainty could be further reduced. Consequently, the risk analyses will

become more reliable and accurate, corresponding more to contents of and frequency of future events. The underlying idea is that if we had sufficient knowledge there would be no uncertainty and thus we would be able to foresee the future, something which of course is not possible (see Laplace's Demon[1]).

Different schools of thoughts meet these challenges differently. Watson [18] presented "The meaning of probability in probabilistic safety analysis" as the degree of belief and proposed it to be a dialectical debate over safety, which is in sharp contrast to those who believes that there is a true underlying risk in the world. The conflict has been very visible between those who view risk as relative frequencies (classical approach) and those who view probabilities (and risk) as a tool to express uncertainty (Bayesian approach). Typical for risk research and practical risk management has been that it is predominately pragmatic in developing risk models, data gathering and risk assessments, neglecting underlying philosophies of science.

2.1.2 The Connotation and Use of the Concept of Safety, Security, Risk and (un)Certainty

Before we discuss the concept of uncertainty, we take a brief look upon the current denotations and etymological basis for some commonly used words in safety management.

The words safe and safety most likely are brought into English via French (adj: *sauf* = without danger) from the Latin verb *salvare*, keep well, live well, save, and the corresponding noun *salvus*, whole, uninjured, intact, not infringed. It is claimed that it is possible to trace the word back to Sanskrit with similar connotations [7].

The original meaning of safety points at an ideal situation or state. This is also parallel to one of the connotations of the word in current English, where safety is defined as the condition of being safe. But according to the Concise Oxford Dictionary, it may also be used as a modifier denoting something designed to prevent injury or damage: a safety barrier. The Macmillan English Dictionary defines safety as the fact that a thing is safe to do or use, but also a place or a situation in which you are protected from danger or harm.

The word security stems from the Latin adjective *securus* which may have a variety of meanings. Ways of translating it into English are free of care, free from worries, happy or cheerful. It has been used in English from late 1500 [7]. At that point of time it replaced the word *sikerte*, which is related to the word *sikkerhet* (Eng. safety) in Scandinavian languages.

Throughout history, the term risk has had varying connotations and uses. Its origin can be traced from the Greek word *rizha*. It denoted something that extended from the trunk of a tree, like a root. Later on in Crete, the beach cliffs "roots" were

[1]Laplace's Demon: a machine that has the capacity to know every detail about the existing world and its intrinsic cause-effect relations, and in addition holding the capabilities to calculate the future based on the preconditions.

given the same designation. It is not clear if these "roots" were seen as a danger for bypassing seafarers and thus were labelled *a risk*, but the word was used in Crete as a nautical expression, a metaphor for "difficult to avoid in the sea". So it may have referred to sub surface (known and unknown) reefs and other types of obstacles that presented a danger if encountered. The word was later absorbed in Latin as the word for *cliff* (*resicum, risicum*), which is the origin of the Italian *risco* and *rischio*, which means danger or risk. In respectively French and Spanish *risque* and *riesgo* were used denoting a similar meaning as in Italian. English borrowed it from Spanish and German borrowed it from Italian. We think it is worth noting that these expressions were spread to different European languages in a time of daring and discovery of new seaways and continents. In the 16th century it seems that risk had a positive connotation. For example in middle-high-German, *Rysigo* (1507) was a technical term for business, with the positive connotation "to dare, to undertake, to hope for economic success".

In the last decades the connotations of risk have become increasingly negative, risk equals dangers and hazards. This can also be seen in several current definitions of risk (examples), and also, as mentioned, in common language connotation. Another explanation for the negative connotation can be found in risks relation to the term *hazard*. Hazard has evolved from the Arabic *al zahr*, which means *the dice*. In Europe playing dice took on a negative connotation as dice games were seen as a possibility for being cheated by con artists. In English as a noun, hazard denotes danger but more interestingly as a verb it denotes *to risk* or *expose to danger*. This is the co-notational connection between the English words risk and danger. If nothing else, this shows that risk and danger are somewhat "related" through the word *al zahr*.

The words certainty and uncertainty appeared in English during the period 1300—late 1400 [7]. It came into English through French from Latin *certanitatem*, which means something that is given or cannot be doubted. Uncertain thus should signify something that can be doubted or discussed.

2.1.3 MSc Students' Perception of Uncertainty When Studying Societal Safety at the University of Stavanger

In the period from 2009–2013, masters students studying societal safety in their second semester have been given a survey. The survey was presented for them when starting up a course in risk management containing the following questions:

1. Describe what you relate to the term uncertainty.
2. Discuss whether uncertainty is related to single persons, or whether groups or entire societies can be connected to uncertainty.
3. Can uncertainty be measured? How would you characterize degrees of uncertainty?
4. If you were asked to express uncertainty related to societal safety, what would you think of?

The response to the survey was organized with two or three students discussing and taking notes, and the responses were then discussed in plenary. There are 59 responses, which are all provided as bullet points or free text sentences.

In general the students listed more than three aspects of the term uncertainty in each questionnaire. The most prominent were firstly, an individual, subjective doubt or lack of knowledge. The second most mentioned characteristic was that uncertainty became related to future outcomes and predictions. Thirdly, uncertainty was closely related to probability, but some also mentioned risk. Furthermore, the students mentioned perceptions like fear and unknowns frequently.

Most of the answers described above were related to individuals, but when students were asked to relate uncertainty to various actors (individuals, groups or society), they usually listed all. However, the answers indicate that it was difficult for them to distinguish between societal systems and societal actors, and similarly for groups the discussions varied from risk exposure to group decisions.

Amongst the students there were a major confusion about measuring uncertainty, they were ambiguous both claiming that it was possible to quantify uncertainty and not. Some equaled uncertainty with probabilities, some claimed that uncertainties could be expressed by risk indices (fatal accident rates, potential loss of life, statistical variation etc.), while many related more qualitative interpretative values (minor, some, much, major) to uncertainty. When the students claimed that lack of knowledge was the premise for uncertainty, they also related psychological characteristics (cognitive, emotional), and even value based considerations (cost-benefit, preferences).

There were no tendencies seen in the students' responses to uncertainty and society. Media communication, disagreement of priorities, disagreement in definitions, vulnerabilities, major accidental events, climate changes, emergency preparedness, diplomacy and politics were some concepts mentioned. Several described complexity as a characteristic that must be given attention when uncertainty should be related to societal safety.

The conclusion to this five-year survey is that even though the students think that uncertainty is an important concept for societal safety, there is no clarity what uncertainty is, how it should be expressed and with what measurement tools it should be considered. Students in 2013 were not different from students in 2009, the responses indicate major confusion about the uncertainty concept. Now we will discuss how uncertainty relates to the scientific stance of risk and safety.

2.1.4 Uncertainty and Its Relation to Risk Theory and Conceptualizations

A crucial issue of any interpretational approach with respect to a scientific theory is the relational elements of the theory on one hand and the elements of the domain of reality for which the theory is designed on the other. This so-called relation of reference cannot be exhaustively addressed within the scope of the scientific discipline concerned. For its proper discussion, genuinely philosophical issues must be taken into account explicitly. One

of the most general demands on a sound philosophical discussion is the distinction between epistemological and ontological statements [4].

Most of the existing scientific and philosophical literature concerning the concept of uncertainty describes uncertainty in epistemological terms. The same can be said about textbooks concerning uncertainty analyses. Much of this literature reflects a normative approach, implicitly stating that the epistemology presented is the most appropriate for solving problems related to uncertainty analyses and decision-making. Seeking the ontological status of uncertainty may yield implications concerning epistemological claims about uncertainty. For example, if uncertainty does not exist as something that is objectively real it cannot be objectively measured as such. If uncertainty can be said to be objectively real it is not a logical or necessary consequence that uncertainty can be objectively measured. Take the core temperature of the sun. The physical state of the core of the sun is an ontological reality, and we can fairly accurately estimate its temperature according to our knowledge of the sun's physical state, but we cannot actually measure the core temperature as such. Further if uncertainty is objectively real; can it be measured objectively?

If you sit in a chair reading, isn't it most likely that you will be in the same situation in ten seconds from now given that you don't choose to do otherwise?. That would be a reasonable assessment. But, it's not absolutely certain that the *status quo* will be upheld. Someone could knock on your door, the phone could ring, the fire alarm might go off, or an airplane might crash into your building. The future is a myriad of different possibilities and thinkable outcomes, but just one is being forever manifest as we go from future, via present to past.

To be able to cope with this myriad of possibilities we need to control or at least attain some sense of control of future events. We do this by making choices. We choose amongst different alternatives of what we would like the future to be. We are constantly being faced with "small" and "big" choices, and we more or less consciously choose amongst the alternatives that present themselves relative to the different choices. The goal of choosing is to satisfy our moment-to-moment or long term needs by choosing the seemingly best (whatever best would mean) alternative at the moment.

The list of "small" and "big" choices that present themselves per day is huge. The choices imply that there are alternate ways of taking action, each action containing the possibility for both losses and gains. Future events, like it could start raining, or I could have an accident on my bike on my way to work, or the land slide occurring at the chemical plant site in 15 years are unknowns at the time of decision. We cannot foresee what all possible and relevant consequences of our choices will be. Nevertheless, we use the concept of risk to help us manage the alternatives that more or less present themselves. We manage risk by using the total sum of all our knowledge, experience and capacity for reasoning to make a decision. The decisions span wide, from the spur of the moment "small" decisions to "big" well considered decisions. We have the "small" minute-to-minute choices we make almost constantly,

and the "large" risks as launching the space shuttle or deciding on building nuclear reactors and so on. The risk connected to the different choices can thus be labelled "small" or "large", but seemingly small decisions can have the potential for large consequences and *vice versa*.

2.2 Contextual Prerequisites for the Uncertainty Concept

Safety and risk are concepts addressing future state of affairs that either is employed for decision making under uncertainty or to characterize an activity's or system's ability to avoid undesired events and outcomes. Thus, the time factor plays a crucial role in order to discuss the ontic status of risk [17].

2.2.1 Time—Past, Present and the Future

2.2.1.1 Uncertainty and the Problem of Tensed Time

The everyday usage and conception of time is its use as a measurement denoting the interval between events. Events are ordered in three main categories consisting of the past, the present and the future. This is sometimes referred to as Newtonian time because this was the scientific interpretation of time until the arrival of the theory of relativity. Measuring time is a human invention, and thus a construction which, according to some philosophers[2] has nothing to do with time per se. In a so called Newtonian time concept the term future is used as an abstraction of all time beyond now. The past being all time before now. Now or present is the term for the transition between past and future.

The philosophical debate on time focuses on the reality of events related to the different tenses. The main controversy has revolved around the status of the future; is the future real? The so called "eternalists" argue that past, present and future events all exist. "Possibilists" hold that past and present events exists. And finally "presentists" argue that only the present is real as such. Discussions concerning the ontic status of time per se have also been a focal point in the philosophical debates. McTaggart [12] formulated what he referred to as an inconsistency in his argument concerning the unreality of time. He argued that since an event E will "move" from future, via present to past as time unfolds, the three states that event E would "go through" would be incompatible, hence the time could not be real as such. McTaggart introduced the terms A-series and B-series of time. These terms have later evolved into separate theories concerning the concept of time. The terms denote two mutually exclusive ways of understanding the concept of time. The A-series is a tensed perception of

[2]See for example Leibniz' correspondence (1715–1716) with Samuel Clarke on Space, Time, and Indiscernible.

time as in past, present and future. Events are seen in relation to one another in the temporal sense of before after and simultaneously. In addition events are categorized as being in the past, the present or the future. The A-series theory is in line with our intuitive understanding of time as tensed, but is not in line with the basic principles of the theory of relativity.

In the B-series theory, time is regarded as tense less. Series of events are related to one another as in the A-series theory, all the way from the Big Bang through present time and into the future. All events in this time series are seen as equally real, thus the ontic status of events that have not yet been manifested are seen as real. This theory is in line with the theory of relativity because tenses are absent.

At least since McTaggart wrote his famous essay "The Unreality of Time", there have been ongoing debates about the ontic status of time and how we should understand the concept of time. This debate has also branched out in debates concerning the ontological status of past, present and future states of the world [6, 9]. An interesting question in the light of the time "controversy" is therefore; does the existence of uncertainty rely on how the concept of time is interpreted? The answer to this is, in part yes. We have previously explained uncertainty as something belonging in the future. Thus we make use of the concept of the future to give meaning to the concept of uncertainty. Consequently one can argue, pending ones philosophical convictions that uncertainty, as something belonging in the future is unreal, hence the "problem" of tensed time and uncertainty. The "problem" being that uncertainty could not be granted a status as objectively real.

However, we think that whether we refer to uncertainty as something of the future or not, is not essential for several reasons. Firstly, it is not absolutely clear that uncertainty is something of the future only, or in part or at all. Secondly, if we referred to the B-series theory and stated that uncertainty is something occurring past a certain point in the time series, the problem would be resolved because tensed time and hence an unreal future would not be of relevance. Thirdly, if the future as such is unreal it does not exclude that logically possible future states of the world can be real, a topic we will discuss later in the article.

What we see as essential to our question concerning the ontic status of uncertainty is whether we in principle can make statements about the outcomes, possible manipulations of situations or actions and to what extent we can render such statements a truth value or a degree of objectivity. Our main conundrum here is that upcoming situations are beyond the present. It is consequently not yet observable, subsequently possible outcomes are in principle unknown. This would imply that our information of what will take place beyond now is in principle zero. In principle anything could happen.

2.2.1.2 What Is Our Knowledge of the Future?

A general view is that if we had adequate knowledge we could reduce the uncertainty and hence predict the future with a greater accuracy. We find this line of thinking inadequate from two positions. First, acquiring adequate knowledge would increasingly

imply a sort of Laplace Demon problem which of course would be impossible to overcome, also philosophically speaking. Second, we think that uncertainty related to the future is of ontic origin not epistemic, because in principle we cannot know the future or future events, thus the uncertainty must be inherently ontic, the uncertainty is a characteristic of the future.

How we can have knowledge of the future is of course dependant on what we mean by the term knowledge. If by knowledge we mean the same as factual knowledge or knowing with certainty it is clear that we cannot know the future in principle, consequently the term "knowledge of the future" would be void of meaning. On such premises one might say that a statement like "the sun will rise tomorrow" would be neither true nor false since we simply cannot know for certain. But if we regard knowledge of the future as an expectation as to what tomorrow will bring or that tomorrow more or less will resemble today, then the term "knowledge of the future" is not meaningless, and the statement "the sun will rise tomorrow" will have a meaningful content. In addition to more linguistic issues related to the knowledge problem there is also the more serious issue of experience. If we did not have any knowledge about possible outcomes of situations or actions, then experience as such would be worthless.

This means that there must be some principles regarding what we could call the stability or regularity of the world, which implies that the world of today will resemble the world of tomorrow. These principles are of course the laws of nature. The laws of nature can at least in part be described as predictive in the sense that they can predict the outcome of experiments (Fig. 2.1). For example the natural law of gravity will predict that if you drop an object in the present it will move in a certain direction with a certain acceleration, and after a specified time the object will be in a specified position in space. But there is a main point to be taken here; it is not absolutely certain that the prediction will be accurate. Something could happen

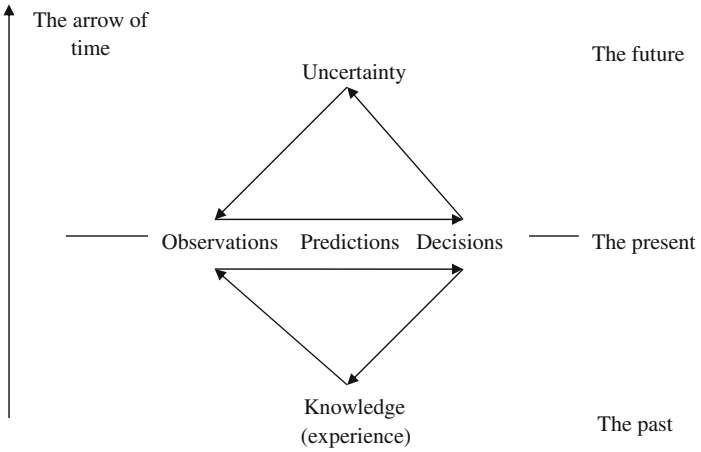

Fig. 2.1 Past, present and future and the concepts of knowledge and uncertainty

which may cause the outcome of the experiment to deviate from the prediction. This is also the core problem with predictions and consequently knowledge of the future; predictions can be made but they are not absolutely certain, they cannot be given any truth value such.

As we see it, this is also the case with the origin of uncertainty as such; we cannot know with certainty what will happen in the future. It is in principle impossible. But we can of course make predictions based on our knowledge about natural laws and other regular phenomena in nature or society. This means that uncertainty is of ontic origin. It is a characteristic of the future and cannot be handled exclusively by increasing our knowledge. Even if there was such a thing as perfect knowledge we still have the problem of imperfect observations and entropic effects thereof. The Danish philosopher Hartnack phrased the problem like this [8]:

> Even in cases where we can have no doubt what so ever about the coming events it is not correct to speak of knowledge. And the reason he gives for this skepticism is well known: All so-called knowledge about the future is based on inductive inferences. We have observed past instances and infer that future instances will be similar. But we can find no reason for assuming that what so far has been repeated will continue to be repeated. Repetition may afford the explanation for our conviction but cannot constitute a logical justification.

2.2.2 System States Through Lenses of Scientific Disciplines

2.2.2.1 Present and logically possible future states of the world

Ontologically, uncertainty is represented by the logically possible future states of the world. In principle, every imaginable logically possible future states of the world represent uncertainty because we do not know what will happen in the future. The future state of the world at time t is uncertain. Therefore we can say that uncertainty is an intrinsic property of the future. The future is, among other things, uncertainty. Thus uncertainty is one of the properties of the future. This is one of the main points in this article. We really do not need any further characterization of the future as uncertain, unknown and so on because the future is per se unobservable and thus unknown. This is the only statement needed to place uncertainty ontologically. Further we must ask what the future contains that is uncertainty.

If we have assessed a specified possible future state of the world, we can monitor and update our assessment as history unfolds. But, if we look at the assessment of a specific state of the world e.g. an offshore installation, it does not imply that after the specified installation were removed that we could measure the total uncertainty of that platform on the grounds that this particular uncertainty has ceased and we now have the complete history. The point being that the uncertainty never ceases. As a result of bringing this installation into production, the logically possible future states of the world will be changed forever, as opposed to not installing the platform that would have rendered different logically possible future states of the world. This means that first you would need a sort of Laplace's demon to render such a calculations possibly

and secondly the calculation would never end. Also observing what happened during the platforms lifespan is not the same as assessing possible future states of the world, thus one would look at uncertainty in retrospect which is quite contrary to what the uncertainty concept is used for.

Ontologically uncertainty is anchored in possible future states of the world, because uncertainty is something of the future in that it is linked to what may or may not happen. But that is not the same as saying that uncertainty *is* some specified logically possible future states of the world per se. The logic for this claim is that no logically possible future state of the world is an unambiguous representation of some specified relation per se. Further, there is no single logically possible future state of the world that would yield a universal agreement of what those possible future states represent. This implies that we do not see any logical argument for separating the logically possible future states of the world and the interpretation of the same because a logically possible future will have no meaning unless we can put meaning into it e.g. by assessing it. Therefore uncertainty is the end product of assessing logically possible future states of the world. The assessment decides whether the logically possible future states present an uncertainty or not, the state of the world in itself does not represent uncertainty implicitly by its own virtue. The different possible states of the world can certainly represent possible dangers and hazards, but then they are just that: possible dangers or hazards. This is because just implying that a logically possible future state of the world represents a danger or a hazard also implies a sort of pre-assessment of those states, which exactly proves our point: who is to decide what is an uncertainty and what is not? This implies that uncertainty is an epistemological category of the present. An implication of this is that if the uncertainty assessor does not exist, uncertainty also does not exist.

2.2.2.2 Present and Logically Possible Future States of the World in Relation to Uncertainty

In principle anything could happen in the future, but intuitively it seems unreasonable that any possible future state of the world should be equally likely. That would mean that life was in a constant state of randomness where nothing could be assessed or predicted and thus leaving the concept of future more or less useless since the chaos today would only be replaced by the "same" chaos of tomorrow. Much to our luck the state of affairs are fairly stable. The earth will most likely continue its orbit around the sun tomorrow, the sun will most likely be producing huge amounts of energy tomorrow and your car will probably be in the garage when you want to go to work tomorrow. The two domains described are contrasting observations of future, but both are equally valid because the unknown future observations exist within the ontological domain and the stable state of affairs observation rests in the epistemological domain. The existence of such a contrasting future in the two domains has obvious implications for the concept of uncertainty on which we will elaborate on below.

We have argued that logically possible future states of the world exist as objectively real. How does this argumentation relate to the ontological status of uncertainty? We argue that it is in principle impossible to claim that any logically future states of the world can be uncertainty free. This would be impossible in principle because the future is unknown. A consequence of this is that logically possible future states of the world represent uncertainty per se, which is our main point in this article.

One could argue that one could theoretically be able to specify *an uncertainty free logical future state of the world* or that such a state is easily imaginable and that this would prove the existence of such an uncertainty free future state. Such an argument would be faulty because it could never be anything but a rhetorical statement, again because the future is unknown in principle. The specification of an uncertainty free future state of the world would just be an epistemological claim since ontologically speaking the future is unknown. However, another and perhaps more eye-catching problem with this relation is that the relation itself is based upon an epistemological claim, namely: no future state of the world can be claimed to be uncertainty free. If our argumentation of an objectively real uncertainty should rest on this would not that implicate an argument of only epistemological value and thus destroying the line of argument? Even if there were no one to observe the future, the future state of such a world could not be uncertainty free because as long as something exists (is objectively real) it will always be exposed to future events, and the consequences of those events cannot in principle be claimed to be uncertainty free (similar as to the previous argumentation). There is a difference between what we in principle do not know and that some future states of the world are more likely to manifest than others. The point being that the future does not become more known or less uncertain because of this difference, this difference is the basis for all claims concerning uncertainty; which again implicitly states that all claims concerning uncertainty is subjective or relative.

If a logically possible future state of the world is being uncertainty, then any thinkable future state of the world would represent an uncertainty, because it would with certainty affect someone in some way (positively or negatively). This implies that the ontological status of uncertainty is totally irrelevant for any epistemological claims about uncertainty, because uncertainty is, ontologically speaking, everything. Certain states of the world can be interpreted as holding a potential for certain consequences by the possible potential effects they represent, but then they would represent just possible effects. Whether we choose to interpret those effects as an uncertainty, is a different matter. This means that uncertainty is an epistemological category only, and this implies that all uncertainty assessments are subjective or relative. This also implies that uncertainty does not have an objective existence per se. What ontologically exist as such are the different logically possible future states of the world and the possible effects they represent. Whether they constitute an uncertainty or not is a different matter.

2.3 Perspectives on Uncertainty in Various Enterprises/Sectors

2.3.1 Health Sector

The concept of risk has no uniform definition in medicine and health care. One of the normative dictionaries merely defines a risk as *a danger or hazard, the probability of suffering harm or other unfavorable outcome* [5].

In epidemiology, which is the part of medicine where the concept of risk has been most thoroughly studied, a quite common definition is *probability of an event during a specified period of time*, where the relevant events e.g. are deaths or appearance of individuals with a defined disease [14–16]. Also in contemporary academic based clinical work (evidence based medicine) this approach seems to be prevailing.

This approach implies that the concept of risk has at least three important properties. Firstly, risk is regarded as something that can be expressed in a quantitative way. Secondly, the calculation of risk emphasizes available historical data of a statistical nature. Thirdly, risk is regarded as a property applicable to populations. Uncertainty is in this context related to aleatory uncertainty and expressed by statistical procedures. The underlying connotation is that there is true risk and probability distributions in which an image of the future is embedded.

The implications of these properties are that risk as a means for prediction of future events to a high degree is an extrapolation of the past. In turn there are principal as well as practical problems connected with judgments of risk on the level of an individual patient.

Thus one may claim that the current concept of risk as described above is more convenient when dealing with populations and public health challenges than with individuals and specific clinical problems.

To deal with practical, clinical problems there has been a lot of work done on medical decision theory, at least since the 1980s. Partly this work has relied on the Bayes' theorem, and partly even has the characteristics of a more fundamental Bayesian approach to the understanding of probability [19]. A typical clinical (diagnostic and treatment) process on individual level carries close similarities with a Bayesian way of thinking. The clinician uses a broad spectrum of general medical knowledge combined with specific information on the single patient to establish a risk picture of the current situation. And this picture is more or less continuously updated when new knowledge appears and hypotheses are tried and proved false.

A short conclusion on the medical approach to risk can be that the clinician, when dealing with individuals behaves according to a Bayesian model, but as a scientist analyzing properties of populations he or she clearly relies on a frequentist's way of thinking.

2.3.2 Aviation/Helicopter Transport

In aviation the concept of risk is closely linked to the concept of safety.

In the guidance material and regulations given in international regulations (ICAO, EASA), the basic idea is to obtain and maintain an acceptable level of safety. In order to achieve that, you need to manage risk. Risk is to be managed with reference to the ALARP principle. Further risk is defined as a combination of frequency and consequence. The management of risk is seen as a typical risk assessment process divided in different stages. You identify potential hazards, you judge the associated risks, and you define an acceptable level and finally you compare the risk to this level to see if further mitigations are necessary. In short, a standard approach based on rational decision processes where politics are obsolete or not considered.

The guidance material or regulations do not mention uncertainty explicitly. Thus implicitly the definition of risk implies that uncertainty is "handled" on at least two levels. Firstly through the frequency approach, where the probability for an event is containing an element of uncertainty in the classical frequency sense, and secondly uncertainty is introduced through the use of the ALARP principle. The ALARP principle, as dictated in the SMM,[3] states that risk is to be reduced to a certain level which should be acceptable. Given a subjective interpretation of risk (and its related consequences) the accepted level of risk is also subjective and thus related to the stakeholder's interpretation of the risk in question and also the subjective interpretation of what is safe enough, i.e. when are we safe (enough)? In relation to the ALARP principle we can see uncertainty in at least two aspects. Uncertainty related to what constitutes a safe (enough) level and uncertainty related to any mitigation put in place to obtain such a level and the uncertainty of the possible effects of those mitigations. We observe that there are different types of uncertainty on different levels, all related to the use of the concept of risk and the use of the ALARP principle.

To have implemented a safety management system is more or less mandatory throughout the world of aviation. Risk as a concept is fairly new in aviation and that can probably explain a lot of the variation in use. Uncertainty as a term is not used at all in any prescriptive or descriptive documentation within aviation. Thus any interpretation of its use must rely on the use of the concept of risk and its interpretation.

2.4 Concluding Remarks

First, we need to address some remarks upon the risk concept. The epistemological claims with regard to risk promote values that are important to humans, such as life and health, environment, and assets. They also focus on preventing negative social values such as social instability, lack of governance or democracy. Being the predominant and acknowledged risk assessor allows for increased power to infer decisions

[3] ICAO Safety Management Manual, 2009.

and influence risk strategies. Risk can therefore become a powerful political tool in areas way beyond the political arenas. The lack of understanding and agreement or lack of focus concerning risk's ontological status fuels the scientific controversies and research on risk. It appears to be expanding in line with the second law of thermodynamics.

We use risk to help us manage decisions and acts based on all our knowledge, experience and capacity for reasoning. It is at best misleading to use the term uncertainty when describing the future. At worst it brings confusion into the picture. There can be epistemological uncertainty related to models and statistical data, but that has nothing to do with the ontological uncertainty concerning risk. It is absolutely certain that we do not know the future. That is not the same as saying that the future is uncertain. The future is unknown, and will remain unknown. Traditional risk analyses, though, can support decisions related to well-known processes, but how this influences uncertainty related to what the future may look like is not clear. Consequently when facing totally new challenges the expectations related to use of risk analyses and the concept of risk should not be exaggerated.

Within the framework of a linear time concept consisting of the past, the present and not least the future, we claim that uncertainty's ontological status exists on various levels. In the present uncertainty is a purely epistemological category, and in the past uncertainty has its meaning related to what has been observed, recognized and comprehended, thus a methodological challenge.

According to how uncertainty is understood and used in risk analyses and risk management, uncertainty can be handled and consequently reduced. The thought is that if we have more knowledge uncertainty could be further reduced, and the uncertainty analyses will be more reliable and accurate (correspond more to the future). That is to say that if we had all knowledge there would be no uncertainty and thus we would know the future, which of course is impossible (see Laplace's Demon). So where goes the limit of our knowledge? Can we have a little knowledge and thus reduce uncertainty with a small amount or can we have a lot of knowledge and thus reduce the uncertainty a lot? No, the uncertainty of the future cannot be handled. What can be handled is uncertainty in methods and measuring (epistemological). But that only contributes to narrowing the quantification; it has nothing to do with the possible future state of the world per se. In the futuristic perspective uncertainty exists and cannot be reduced.

References

1. Althaus CE (2005) A disciplinary perspective on the epistemological status of risk. Risk Anal 25(3):567–588
2. Aven T (2010) Misconceptions of risk. Wiley, Chichester
3. Aven T (2014) Risk, surprises and black swans. Routledge, New York
4. Castell L, Ischebeck O (eds) (2003) Time, quantum and information. Springer, Berlin
5. Dorland WAN (2000) Dorland's illustrated medical dictionary. Saunders, Philadelphia

6. Hamlyn DW (1984) Metaphysics. Cambridge University Press, Cambridge
7. Harper D (2015) Online etymology dictionary
8. Hartnack J (1959) Knowledge of the future. Philos Stud 10(6):89–96
9. Jacquette D (2002) Ontology. McGill-Quenne University Press, Montreal
10. Jaeger CC, Webler T, Rosa EA, Renn O (2001) Risk, uncertainty, and rational action. Earthscan, London
11. Lipshitz R, Strauss O (1997) Coping with uncertainty: a naturalistic decision-making analysis. Organ Behav Hum Decis Process 69(2):149–163
12. McTaggart JE (1908) The unreality of time. Mind: Q Rev Psychol Philos 17:456–473
13. Morgan MG, Henrion M, Small M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York
14. Porta M (2008) A dictionary of epidemiology. Oxford University Press, Oxford
15. Rothman KJ, Greenland S, Lash TL (2008) Modern epidemiology. Lippincott Williams & Wilkins, Philadelphia
16. Saracci R (2010) Epidemiology: a very short introduction. Oxford University Press, Oxford
17. Solberg Ø, Njå O (2012) Reflections on the ontological status of risk. J Risk Res 15(9):1201–1215
18. Watson SR (1994) The meaning of probability in probabilistic safety analysis. Reliab Eng Syst Saf 45(3):261–269
19. Wulff HR (1976) Rational diagnosis and treatment. Blackwell Scientific, Oxford

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 3
A Conceptual Foundation for Assessing and Managing Risk, Surprises and Black Swans

Terje Aven

Abstract This chapter presents and discusses some recent advances in the risk field, linked to the conceptualisation of risk and specifically addressing unforeseen events, surprises and so-called black swans. It shows how the traditional probability-based perspectives on risk are extended to broader ways of thinking about risk, which give due attention to the uncertainties and also draw on ideas from *inter alia* the quality discourse and organisational learning (collective mindfulness and its five principles: preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience and deference to expertise). The main aims of the chapter are to point to this thinking and provide some reflections on how to use it for further developing the risk assessment and risk management fields.

Keywords Uncertainty · Risk perspectives · Risk concept · Risk assessment · Surprises · Black swans

3.1 Introduction

In recent years there has been an increased focus on uncertainties in relation to risk assessment and risk management. The traditional probabilistic approach to risk assessment and management is criticised for its narrowness in the way it looks at risk and how it copes with uncertainties (see e.g. [15, 52]). In response to the critique, several alternatives have been suggested. These can be grouped as follows:

1. Replacing the probabilistic approach with other quantitative approaches, for example based on interval probabilities (typically supported by possibility theory or evidence theory) [8].

The copyright to third-party material quoted in this chapter is property of its respective owners and is not included in the Creative Commons license for this chapter.

T. Aven (✉)
University of Stavanger, Stavanger, Norway
e-mail: terje.aven@uis.no

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_3

2. Balancing alternative approaches, in particular the probabilistic approach and approaches that are effective in meeting hazards/threats, surprises and the unforeseen. For short, we refer to the latter approaches as "robust approaches".
3. Rejecting the probabilistic approach to risk assessment and risk management, and instead relying on robust approaches.

Robust approaches cover cautionary measures such as designing for flexibility (meaning that it is possible to utilise a new situation and adapt to changes); implementing safety barriers; improving the performance of barriers by using redundancy, maintenance, testing etc.; and applying quality control/assurance. It also covers concepts like resilience engineering, which is concerned with finding ways to enhance the ability of organisations to be resilient in the sense that they recognise, adapt to and absorb variations, changes, disturbances, disruptions and surprises [35]. We may also include the concept of antifragility [50]. According to Taleb, the antonym of fragile is not robustness and resilience, but "please mishandle" or "please handle carelessly"—to quote an illustration from Taleb when referring to sending a package full of glasses by post. The antifragile is seen as a blueprint for living in a "black swan world", the key being to love randomness, variation and uncertainty to some degree, and thus also errors. As our bodies and minds need stressors to be in top shape and improve, so do other activities and systems [9, 12].

In practice, alternative (2) normally applies. If we look at high risk industries such as nuclear and oil & gas, we find a mixture of the probabilistic and the robust approaches. It is acknowledged that the probabilistic approaches have limitations in managing risk, surprises and the unforeseen, and need to be supplemented by robust approaches. For societal safety and security contexts, we see this duality as even stronger, for example in relation to terrorism risk. Here probabilistic risk assessments are hardly used. The information provided by assigned attack probabilities is small in most cases [7, 21].

To confront risk and uncertainties, robust thinking is needed, to a varying degree depending on the situation. This is also true when alternative quantitative approaches (approach 1) are adopted. Using interval probabilities in place of specific probability numbers can allow for more balanced judgements reflecting what is known and what is not, but, as thoroughly discussed in [8], the interval probability approach is not easily implemented in practice—there are many challenges—and most importantly, it cannot replace the need for robust approaches. No analytical quantitative approach, probabilistic or not, can make robust arrangements and measures superfluous, as they will always be subject to some limitations. They can provide useful decision support, but not prescribe what to do.

This acknowledgement implies the need for a way of thinking about risk that sees beyond the probabilistic perspective. In engineering applications, risk is commonly seen as expected values (probability multiplied with loss) or as the combination of probability and consequences (loss) [5, 19, 38]. However, such a perspective is too narrow to adequately capture the balanced setting of the approaches of category (2), and in recent years we have seen many attempts to conceptualise risk to meet this broader scene. One example is the new definition of risk adopted by the ISO (2009): "risk is the effect of uncertainties on objectives". A key aspect with this def-

inition of risk is that uncertainty replaces probability. This may seem to be a rather minor change, but it has important implications. It will be thoroughly discussed in the following when reviewing recent developments in the risk field, paying special attention to the perspectives that can serve the assessment and management approaches (2) mentioned above. The review, which will be partly based on [5], shows that there are many frameworks and approaches suitable for (2), which give further substance and precision to the understanding of risk compared to, for example, the ISO risk definition. From this review we will discuss in more detail how surprises and the unforeseen (black swans) are coped with in these frameworks and approaches. The main purpose of the chapter is to provide new insights on the link between these concepts—risk, surprises and the unforeseen—and to give guidance on how to adequately think in this respect, first when it comes to conceptualisation, then next when it comes to the way we should assess and manage risk, surprises and the unforeseen.

3.2 Risk Perspectives, Brief Review of Historical and Recent Development Trends

The origin of the term "risk" is disputed in the literature, see e.g. [1] and [5], but there seems to be broad agreement between researchers that De Moivre's 1711 definition is one of the first formal definitions of risk used in a risk analysis context [25]. De Moivre defines the risk of losing any sum to be the product of the sum adventured multiplied by the probability of the loss, i.e. risk is defined as the expected loss. This definition is still being used in many contexts despite the strong arguments against it; see summary of arguments in [5], Appendix A. About 200 years later, expected loss is replaced by probability: risk is the chance of damage or loss [33]. This perspective is further developed to reflect the magnitude of the losses and consequences, and at the beginning of the 1980s the so-called triplet definition of risk was the dominating understanding of the risk concept, at least in engineering contexts, which covers events/hazards, the consequences of these and associated probabilities [38].

In the social sciences this probability-based approach to risk is challenged and alternatives are proposed, one of the most popular being Rosa's definition expressing risk as a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain [30, 53].

Many researchers from the social sciences refer to risk but do not distinguish between risk per se and how risk is managed. One illustrating example is the German sociologist Beck, who states that "Risk may be defined as a systematic way of dealing with hazards and insecurities introduced by modernization itself" [17, p. 21]. It represents a way of looking at risk which is in conflict with most other perspectives on risk, and it is difficult to justify. To use the words of [22, p. 151], in their investigation of Beck's work on risk:

> It is hard to think of a less adequate definition: risk is not a way of dealing with things... Beck's definition would make it impossible to ask: How are we responding to this risk?, as the response and the risk would be the same thing. Secondly, risk should not be so defined that it applies only to 'modernization', for there were of course risks before industrial society.

We also see that the concept of risk and risk perception is mixed. The perception dimension includes personal feelings and affections (for example dread) about the possible events, the consequences of these events and about the uncertainties and probabilities, and even judgements about risk acceptability. According to cultural theory and constructivism, *risk is the same as risk perception* [29, 37, 65]. Beck [17, p. 55] states that "because risks are risks in knowledge, perceptions of risks and risk are not different things, but one and the same". Also this way of thinking clashes with most professional and scientific risk perspectives, as they seek to distinguish between what is risk and what are feelings and affections, and value statements about risk (for example, what is acceptable risk).

In the economic environment, risk has also been associated with uncertainty, for example by Hardy as early as 1923 [32]. He states that risk is uncertainty in regard to cost, loss or damage. We also see this perspective today, and typically the uncertainty is expressed by the variance or standard deviation. The case that the risk is objective uncertainty is also commonly referred to today. This perspective is to a large extent based on Frank Knight's work from 1921 and his conceptualisation, where risk is used for the case that an objective probability distribution can be obtained (and uncertainty otherwise) [42]. This nomenclature has strongly influenced the risk area, and in particular the economic risk field. Referring to risk only when we have objective probability distributions would mean that we exclude the risk concept from most real-life situations, and this terminology is therefore avoided by most risk analysts and researchers (see relevant references and discussion in [5]). Clearly, if one adopts the subjective or Bayesian perspective on probability, Knight's definition of risk becomes empty. No objective probabilities exist.

The risk = uncertainty perspective is typically based on the assumption that the expected value is the point of reference and that it is known or fixed. The uncertainty is seen in relation to, for example, a historical average value for similar investments. Risk captures the deviation and surprise dimension compared to this level. Without such a reference level, the "risk = uncertainty thesis" does not work. Uncertainty seen in isolation from the consequences and the severity of the consequences cannot be used as a general definition of risk. Large uncertainties need attention only if the potential outcomes are large/severe in some respect; see example in [5].

This leads us to risk perspectives highlighting both losses/consequences and uncertainties. For short we refer to these as the (C, U) perspectives, C denoting the consequences of the activity and U the uncertainties (what will C be) [5, 13]. The ISO (2009) definition, where risk is being defined by the effect of uncertainty on objectives, can be seen as a special case of this definition, with the consequences linked to objectives [5]. The Petroleum Safety Authority Norway has implemented a new definition of risk in line with the (C, U) definition: For the area of health, working environment and safety, risk means the consequences of the activities, with associated uncertainties.[1]

[1] Risk definition of the Norwegian Petroleum Safety Authority, accessed 7th October 2015, http://www.psa.no/risk-and-risk-management/category897.html.

The consequences C can be seen in relation to objectives or some other references, for example some planned numbers, and the focus is normally on negative, undesirable consequences. There is always at least one outcome that is considered as negative or undesirable. The consequences are with respect to something that humans value (including life, health, environment, economic assets).

The consequences also cover events, such as a gas leakage or other hazardous events. The knowledge dimension enters the scene when we try to measure or describe risk. This is performed by specifying the consequences C and using a description (measure) of uncertainty Q (which could be probability or any other measure—measure here interpreted in a wide sense). Specifying the consequences means to identify a set of events/quantities of interest C' that characterise the consequences C. An example of C' is the number of fatalities. Depending on which principles we lay down for specifying C and the choice of Q, we obtain different perspectives on how to describe/measure risk. As a general description of risk, we can write

Risk description $= (C', Q, K)$, where K is the background knowledge on which C' and the assignment Q are based.

Adopting this risk perspective, it is argued in [3] that safety is the antonym of risk, and that the analysts conclude on safety being high or low, and being safe, by reference to the risk description (C', Q, K).

We see that such a way of understanding and describing risk allows for all types of uncertainty representations, and it could consequently serve as a basis of a unified perspective on uncertainties in a risk assessment context. The most common uncertainty representation is probability, but, as mentioned in Sect. 3.1, there are also others (including those based on interval probability, possibility theory and evidence theory) [15]. In all forms of risk descriptions, the knowledge dimension (data, information, justified beliefs) and the strength of this knowledge need to be seen as an integral part.

The risk descriptions are often based on modelling, for example using probability models and frequentist probabilities. Probability models and frequentist probabilities do not exist in general: they are model concepts that are meaningful only in some situations of repeatability.

A basic way of categorising risk is related to the distinction between conceptualisations of risk that see risk as an objective property of the world and conceptualisations that are based on judgements and knowledge of a person (i.e. are epistemological). We remember the well-known phrases used by Immanuel Kant (1724–1804), "Das Ding an sich" and "Das Ding fur mich". Risk (and probability) can be viewed as both an "an sich" property of the world and a "fur mich" concept; see discussions in [4, 60].

In their famous 1981 paper [38], Kaplan and Garrick also refer to risk as qualitatively defined as "uncertainties + damage", which can be seen as a version of the (C, U) type of definitions. However, the authors did not develop a theory for this perspective as shown above for the (C, U) – (C', Q, K) approach (this theory is based on early work by, for example, [2, 11]).

In this chapter, we are especially concerned about how the risk perspective supports the integration of concepts like unforeseen events, surprises and black swans.

In a recent work a new perspective on how to conceptualise, assess and manage risk and the unforeseen, covering surprises and black swans, is presented [12]. This work builds on the (C, U) risk perspective and draws on ideas from the quality discourse and the use of the concept of collective mindfulness as interpreted in studies of High Reliability Organisation (HRO). It also provides a suggestion for a classification of black swan type of events. We will look more closely into this perspective in the following section.

3.3 Risk, Surprises and Black Swans

An event is commonly considered a surprise when it occurs unexpectedly and also runs counter to accepted knowledge [31]. There exist, however, many other definitions, for example: a surprising event may be regarded as one whose occurrence was not anticipated, or which has been allocated such a low probability that the possibility of its occurrence was effectively discounted [40, p. 69]. The literature includes many taxonomies for classifying surprises. Examples include the dichotomies between known (imaginable) surprises and unknown surprises, between unanticipated surprises and anticipated surprises, and between the unintended, imaginable and anticipated [31, pp. 37–41]. An imaginable surprise occurs when the event type is known but its occurrence was considered highly unlikely. If we know that something is going to happen, but not when and in what form, it is referred to as an anticipated surprise. As noted by [31, p. 40], a surprise cannot be registered in any meaningful way without an "expectation" in some sense, to create a deviation.

In 2007, Nassim Taleb defined and popularised the concept of black swans in his book, *The Black Swan* [49]. Taleb defines a black swan to be an event with the following three attributes: firstly, it is an outlier, as it lies outside the realm of regular expectations, because nothing in the past can convincingly point to its possibility. Secondly, it carries an extreme impact. Thirdly, in spite of its outlier status, human nature makes us concoct explanations for its occurrence after the fact, making it explainable and predictable.

Several alternative definitions have been suggested, including the one recommended by [8]: a black swan is to be seen as a surprising extreme event relative to the present knowledge/beliefs. Hence the concept always has to be viewed in relation to whose knowledge/beliefs we are talking about, and at what time. Aven and Krohn [12] build on this definition and distinguish basically between three types of such events:

(a) Unknown unknowns: events that were completely unknown to the scientific environment (unknown events to us, unknown to others);
(b) Unknown knowns: events not on the list of known events from the perspective of those who carried out a risk analysis (or another stakeholder), but known to others (unknown events to us, known to others);
(c) Events on the list of known events in the risk analysis but judged to have a negligible probability of occurrence, and thus not believed to occur.

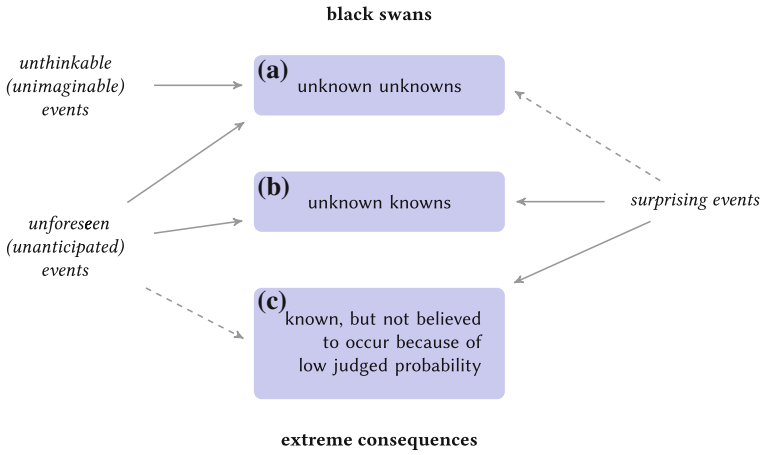

Fig. 3.1 Schematic illustration of the concepts *black swan*, unknown unknowns, unforeseen events, surprising events and unthinkable events, based on the ideas presented by [12] and presented in [8]

The term "black swan" is used to express any of these types of events, tacitly assuming that it carries an extreme impact. See Fig. 3.1, which links terms such as black swans, surprising events and unforeseen events.

The first category (a) of black swan type of events is the extreme—the event is unthinkable—the type of event is unknown to the scientific community, for example a new type of virus. In activities about which there is considerable knowledge, such unknown unknowns are likely to be rarer than in cases of severe or deep uncertainties.

The second type of black swan (b) is events that are not captured by the relevant risk assessments, either because the analysts do not know them, or they have not made a sufficiently thorough consideration—the events are known to others. If the event then occurs, it was not foreseen. If a more thorough risk analysis had been conducted, some of these events could have been identified. The third category of black swans is events that occur despite the fact that the probability of occurrence is judged to be negligible. The events are known, but considered so unlikely that they are ignored; they are not believed to occur and cautionary measures are not implemented. Here is an example: an underwater volcano eruption occurs in the Atlantic Sea resulting in a tsunami affecting Norway. The event is on the list of risk sources and hazards but then removed as the probability is judged to be negligible. The occurrence will come as a surprise. The tsunami that damaged the Fukushima nuclear plant was similarly removed from the relevant risk lists due to the judgment of negligible probability.

A (unanticipated) surprising event (with severe impacts) is thus a black swan according to this logic. An event of category (b) and (c) will obviously come as a surprise, but it is not so obvious when we talk about category (a) unknown unknowns—whence the dotted arrow in Fig. 3.1. Considering an activity with deep uncertainties about the type of events that will occur and the impact they will generate, we may be completely free of "expectations" for what is coming. Hence, it may be questioned whether an unknown unknown is in fact coming as a surprise in such a situation.

Similarly, we may problematise what is an unforeseen event. If an event occurs which was judged to have negligible probability, was it then foreseen? Yes, in the sense that the fact that it could happen was anticipated, but not in the sense that it was not considered likely.

Think of a container of fluid; normally it is filled with water and daily people drink from it. One day Ole drinks fluid from the container and it turns out to be of a toxic type. We refer to it as black swan, a surprise in relation to his knowledge/beliefs (assuming that it carries a serious impact). It need not be a new phenomenon we are witnessing, or an unknown unknown to be labelled a black swan. In retrospect, we explain the incident easily.

A risk analysis could have identified such an event; nevertheless, it may be surprising for some people (Ole), in relation to their beliefs/knowledge. These are the type of events we are concerned about. Let us modify the example a little. Suppose a risk analysis has identified various types of toxic fluids that could fill the container in special situations, but excludes a dangerous form because of a set of physical arguments. But then, this scenario happens. The event was possible despite the fact that it was considered impossible (extremely unlikely) by the analysts. The real-life conditions were not the same as those that were the basis for the risk analysis, and the event came as a surprise even for the risk analysts. In retrospect it was, however, easily explained.

In relation to the classifications (a)–(c) above, the event is classified as belonging to category (c). For the first case (Ole), the basis is the beliefs that this person has, and it can then be placed under (b) or possibly (c).

Strictly speaking, it would make sense to say that an unthinkable event is an unknown unknown. We may, however, also argue differently. Viewed from a risk assessment point of view, an event that belongs to category (b) may also be judged as unthinkable provided a thorough analysis has been performed to uncover all relevant events. The question is: unthinkable for whom?

It is common to refer to the term "unexpected" when characterising surprises and black swan type of events. However, this term is problematic to use. Consider the following example. An event has three possible outcomes, 0, 50 and 100, with associated probabilities 0.25, 0.50 and 0.25, respectively. Hence the expected value in a statistical sense (the centre of gravity of the probability distribution), is equal to 50. The values 0 and 100 can thus be seen as unexpected; however, the probability of one of these outcomes occurring is 50 %, which cannot be viewed as surprising. Clearly for this term "unexpected" to make sense, we need to interpret it in relation to the probability distribution. What is "unexpected" needs to be understood more as outcomes not belonging to a sufficiently broad uncertainty interval $[a, b]$, such that the probability of the quantity of interest not being covered by this interval is small, say less than 5 %.

3.4 Assessing and Managing Surprising Events and Black Swans

It is a common perception that the challenge in dealing with surprises and black swans lies in the fact that they are beyond the sphere of probability and risk (e.g. [31]). However, in line with the new (C, U) risk perspectives, the issue of confronting such events is at the core of risk management. A suitable conceptualisation for this has been established, summarised in the previous section; now the challenge is to develop adequate types of risk assessments and management policies. Let us first reflect on how this can be done for risk assessment.

3.4.1 Assessment

Many types of traditional risk assessment methods address the issue of what can happen, for example HAZOP, HazId, fault tree and event tree analysis [66]. Using these methods, hazardous events and scenarios are identified, but as risk is an issue, uncertainties and likelihood related to these events and scenarios are also addressed, in a qualitative or quantitative way. Based on considerations of probability, some events and scenarios can be judged to pose a negligible risk and not be followed up any further. With an increased focus on surprises and black swan types of events and scenarios, we need to reconsider the use of these methods. We have to challenge the premises that the analyses are based on, the assumptions that the probability judgements are relying on, etc., as for example indicated in [6, 12]. It is beyond the scope of this chapter to present specific methods for how to carry out such extended analyses; rather, the aim is to point to some key challenges and provide some preliminary reflections on possible routes for the research and development within these areas.

Kaplan et al. [39] provide some ideas for new tools in this direction, the so-called anticipatory failure determination (AFD). This method is an application of I-TRIZ, a form of the Russian-developed Theory of Inventive Problem Solving. This method is particularly suitable for scenarios involving human error, sabotage, terrorism, and the like. The relevance of the TRIZ methodology to risk analysis is rooted in the fact that revealing and identifying failure scenarios is fundamentally a creative act, yet it must be carried out systematically, exhaustively, and with diligence [39]. Traditional failure analysis addresses the question, "How did this failure happen?" or "How can this failure happen?". The AFD and TRIZ go one step further and pose the question, "If I wanted to create this particular failure, how could I do it?" The power of the technique comes from the process of deliberately "inventing" failure events and scenarios [48]. See [8] for a simple application example of this method.

These methods can be supported by different types of analysis frameworks, for example, Actor Network Theory (ANT) [44, 48]. ANT seeks to understand the dynamics of the system by following the actors – it asks how the world looks through

the eyes of the actor doing the work. As highlighted by [27, p. 1630] and [48], through this approach issues emerge pertaining to the roles that tools and other artefacts (actors) play in the actor-network in the accomplishments of their tasks.

TRIZ provides one type of creative thinking; there are many others, as discussed by for example [57]. There is obviously a potential for further developments of analysis methods for black swan types of events using some of these. The first step would be to assess the suitability of the various techniques for such a purpose.

The area of scenario analysis can add valuable input to the risk assessments. Here scenarios are developed describing potential future conditions and events, using various techniques [23]. There is no search for completeness and characterisations of the uncertainties and risks, as in traditional risk analysis, but in the case of large uncertainties and a lack of accurate prediction models, the generation of such scenarios may provide useful insights about what could happen and possible black swans. The deductive (anticipatory, backwards) scenarios are of particular importance in this respect, where we start from a future imagined event/state of the total system and question what is needed for this to occur. System thinking, which is characterised by seeing wholes and interconnections, is critical if we are to identify black swans, as for example highlighted by many scholars of accident analysis, organisational theory and the quality discourse [28, 58]. Using for example techniques such as event trees to reveal scenarios has strong limitations, as the analysis is based on linear inductive thinking [35, 46].

The risk description of an activity could be strongly influenced by the available signals and warnings, and hence we need to question to what degree the risk assessments are able to reflect these signals and warnings, in particular attitudes such as awareness and mindfulness in relation to these. A number of techniques exists which can contribute to enhance such attitudes, for example the use of the Johari Window concept [63] and ideas from the collective mindfulness concept used for High Reliability Organizations [61]. Further developments are, however, required to better link signals and warnings, with the associated awareness and mindfulness concepts, to risk. Some initial reflections on the topic are provided by [12, 41].

Another method that may be useful in revealing black swans is *red teaming*, which serves as a devil's advocate, offering alternative interpretations and challenging established thinking [48]. For example, "businesses use red teams to simulate the competition; government organisations use red teams as 'hackers' to test the security of information stored on computers or transmitted through networks; the military uses red teams to address and anticipate enemy courses of action" [16, p. 136]. Red teaming challenges assumptions, generalisations, pictures or images that influence how we understand the world and how we take action, i.e. our mental models [54]. [26, p. 39] argues that Murphy's Law is wrong: everything that can go wrong usually goes right, and then we draw the wrong conclusions. Red teaming can help us to see why, by pointing to alternative scenarios and outcomes.

Probabilities and interval probabilities may be used to express uncertainties and degrees of belief, but equally important is the knowledge these are based on. Crude probability judgements can indicate that some events are unlikely, but such statements need to be supplemented with assessments of strength of knowledge that the

judgements are based on and in particular the assumptions that they are based on. Ideas for how this can be conducted is outlined in [5] and [6], for example using the concept of assumption deviation risk.

3.4.2 Risk Management

To manage risk, three major categories of measures/strategies are commonly referred to [14, 52]: risk-based approaches (use of risk analysis), cautionary and precautionary approaches (including strategies based on robustness and resilience), and discursive strategies. The risk-based approach can be used alone only in cases when the knowledge is very strong and the uncertainties small. In most situations, all three strategies are required. The challenge is to find the adequate balance between these approaches and strategies, in particular between the first two. When the stakes are high and the uncertainties large, we obviously need to highlight robust and resilient solutions and arrangements to be prepared in case some extreme unforeseen events should occur. Potential surprises and black swans call for robustness and resilience, and antifragility, as argued for by [50], see also Sect. 3.1 and [10].

Adaptive risk management also needs to be mentioned here. This is a type of management which seeks to treat risk by considering a set of alternatives and dynamically tracking these to gain relevant information and knowledge about the effects of different courses of action [34, 64]. Such a strategy is especially attractive in the case of large/deep uncertainties, as discussed in [24]. This reference also points to several other risk management tools that can be used to confront deep uncertainties in a risk context, including learning what to do by well-designed and analysed trial and error. See [20] for an offshore application of adaptive risk management in a (C, U) risk perspective.

The risk-based approaches incorporate risk assessments but need to be extended and have a broader scope than the standard probabilistic analysis commonly seen in textbooks and practice today, as indicated in the previous section. A focus on knowledge building, transfer of experience and learning, represents an important means to reduce the risk related to surprises and black swans, by obtaining improved understanding of relevant systems and activities, models and the ability to predict what is coming. To provide a suitable foundation for such improvements, we need a platform that incorporates adequate concepts, assessment and management principles and methods. This is a research issue, and some ideas for how such a platform can be defined are presented by [12]. The idea is to integrate the conceptual framework of the (C, U) risk perspective as outlined in Sect. 3.2, with associated assessment and management principles and methods, and add theories and practical insights from other fields specifically addressing the knowledge dimension and the black swans. In [12], two areas are given main attention, first the collective mindfulness concept with its five principles: preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience and deference to expertise. These elements represent criteria for checking the soundness of the understanding and treatment

of the risk at various stages of the accident scenario, including the planning phase of the actual operation. There is a vast amount of literature (see e.g. [36, 45, 61, 62]) demonstrating that these five principles explain High Reliability Organizations (HROs) well and that the collective mindfulness concept thus can be used as an effective instrument for managing risks, the unforeseen and potential surprises.

In addition, the use of ideas from the quality discourse, with its link to the concepts of "common-cause variation" and "special-cause variation", and the continuous focus on learning and improvements, are added. Common causes capture "normal" system variation, whereas the special causes are linked to the unusual variation and the surprises (black swans) [18, 28, 55, 56].

There are a number of other risk management approaches, based on different ideas and traditions, which could be added to those briefly mentioned above. See for example [43, 46, 47]. However, for the purpose of the present chapter, the above review suffices, as it provides a basis for discussing how the integrated framework and thinking outlined above (summarised in [12]) can be used in a practical risk management context. The main uses of the framework are:

1. As a general guideline for designing methods and tools for understanding, assessing, managing and communicating risk and safety
2. As a means for evaluating (also covering suggestions for improvement) the quality ("goodness") of various risk management activities, such as use of various types of risk assessment, barrier principles, risk management strategies (for example use of the ALARP[2] criterion)
3. As a supplement to standard accident investigation procedures, by drawing attention to critical issues such as the violation of the five collective mindfulness principles, and issues related to knowledge, uncertainties, variation, the unforeseen, surprises and black swans, from different perspectives (individuals and groups) and points in time
4. As a means for evaluating the quality ("goodness") of a concrete risk assessment being conducted
5. As a means for evaluating the quality ("goodness") of the risk understanding in relation to critical operations, reflecting the ability to understand the total system, relevant knowledge, transfer of experience and learning.

To illustrate the use, let us look into some examples from the petroleum industry.

As a first example, let us consider the guidelines from the Norwegian Petroleum Safety Authority (PSA-N) for barrier management principles [51]. With respect to concepts, principles and main thinking for the understanding, assessment and management of risk and safety, it can be concluded that the document is very much in line with the integrated framework presented here. This applies, for example, to the meaning of the concept of risk. A further analysis of the five collective mindfulness principles and aspects highlighted by the quality discourse reveals two main challenges. The first one relates to the reluctance to simplify criterion of the collective mindfulness concept. Although the barrier principles highlight the need for a total

[2] ALARP: As Low As Reasonably Practicable.

system view with a focus on total barrier functions, the strong emphasis on the specification of detailed barrier element performance requirements may lead to difficulties in practice. Meeting the barrier element performance requirements may give the false perception that the risks are low and the barrier functions fulfilled. As we know, the connections between barrier element performance, risk and satisfying barrier functions are often unclear. Good performance numbers for the detailed barrier elements are no guarantee of safety. Holistic thinking is important, particularly for being able to identify black swans as well as for ensuring robustness and resilience. From a practical point of view, we may find that both the industry and the agency, through their auditing, are happy with a regime that highlights the barrier element as it is simple and easily followed up, compared to broader more judgemental assessments of barrier function performance and risks. The PSA-N is aware of this challenge, and it will be interesting to see how the implementation process will proceed.

The second challenge of these barrier management principles relates to management by objectives and the compliance focus, and is linked to the discussion above concerning the reluctance to simplify. The approach recommended by the PSA-N document has a strong emphasis on formulating, assigning and satisfying performance requirements, which can easily lead to an excessively strong focus on meeting requirements rather than on identifying the overall best solutions and measures. Experience has shown that such an approach represents a serious challenge – the compliance regime prevails, much more than the improvement processes which are always highlighted in theory but often fail to be given priority when in competition with the convenience and practical attractiveness of compliance procedures.

The second example of applications 1.–5. relates to the use of job safety analysis in an industrial setting. In a job safety analysis, hazards linked to the various subtasks are identified, and risk is described by a risk matrix covering consequence and probability. Based on this risk description, an overall judgement about safety and risk acceptability is made. An evaluation of this method (Use 2), quickly reveals that the way risk is described is based on a probability-based thinking on risk, and, as discussed in Sect. 3.2, this perspective can be criticised by not properly reflecting the strength of knowledge that the assessments are based on, uncertainties, surprises and black swans. It is beyond the scope of the present chapter to provide a detailed solution to this challenge, but some ideas have already been pointed to above, in the last paragraph of Sect. 3.4.1, by incorporating judgements of the strength of knowledge that the probability-based judgements are based on, and making assessments of the so-called assumption deviation risk. Specific related black swan type of analysis should also be considered, some type of red-team analysis as mentioned in Sect. 3.4.1, for example using the ideas outlined in [6]:

> First make a list of all types of risk events having low risk as scored by the three dimensions: assigned probability, consequences, and strength of knowledge. Then carry out a review of all possible arguments and evidence for the occurrence of these events, for example by identifying historical events and experts' judgements not in line with common beliefs. To conduct these assessments, experts are needed that are not members of the core group of analysts conducting the more standard parts of risk assessment. The idea is to allow for and stimulate different views and perspectives, in order to break free from the prevailing beliefs and obtain creative processes.

Further examples can be found in [8]. This leads us to the third example, linked to the fourth use of the integrated framework. A risk analysis team has conducted a risk assessment in relation to a critical maintenance operation, a job safety analysis. To evaluate the quality of this analysis, we apply the principles of the integrated framework, presuming that the intention now is to conduct it in line with the principles just outlined. A check is made that the analysis covers what can go wrong, causes, barriers, consequences and likelihood, as well as issues linked to uncertainties, knowledge, surprises and black swans. A key point is that the analysis is clear on what are the key assumptions and beliefs that the judgements are based on, and that these assumptions and beliefs are challenged as discussed in an appropriate way. Key control questions relate to the data and information supporting the assessment, and the ability of the analysts to provide a meaningful presentation of the results, which is clear on what the analysis provides and what the limitations are. In the concrete example studied, it turned out that the results relied on a key assumption which was not questioned in the uncertainty analysis, despite the fact that the review showed that some critical questions to it had been raised in some other contexts. These questions were, however, not known to the analysis team first performing the risk analysis.

This critical maintenance operation is also considered in the example adopted for the fifth use. We would like here to apply the framework to evaluate the risk understanding of personnel that are to carry out the work. The focus is on the potential hazards, their causes and consequences, including barriers, as well as the likelihood. Special attention is paid to signals and warnings that can give an indication of the occurrence of some more severe hazardous situations, and how to sense these signals and warnings and make adequate adjustments. The degree to which the insights provided by the relevant risk assessments are known and understood is questioned. This relates to all the issues mentioned above (what can go wrong, causes, barriers, consequences and likelihood), but also covers issues linked to uncertainties, knowledge, surprises and black swans (if such issues have not been addressed by the risk assessment, a separate analysis of such issues needs to be added to be in line with the recommended thinking). A key issue is to reveal what are the main assumptions and beliefs that the judgements are based on, and how to cope with surprises relative to these assumptions and beliefs. In the specific case considered here, the fourth collective mindfulness principle was pointed to as a weakness, as the issue of robustness and resilience had not been properly thought through. The current standard in the company for how to carry out such operations, fails to stress the importance of commitment to resilience. Moreover, it was identified that key personnel had a lack of understanding of the system, and that important experience from some recent operations had not been captured by the risk assessment and was not known to all the personnel involved in carrying out the critical operation.

For an example linked to accident investigation (Use 3.), see [59, p. 82]. See [6, 12] for some other examples of Use 2., linked to risk acceptance criteria and the implementation of the risk reduction principle ALARP.

3.5 Conclusions

We have reviewed some recent advances in the risk field, linked to the conceptualisation of risk and specifically addressing unforeseen events, surprises and so-called black swans. Some ideas for new types of risk analyses were briefly outlined before an integrated framework was presented which extends the traditional probability-based perspectives on risk to broader ways of thinking about risk, which give due attention to the uncertainties and also draw on ideas from the quality discourse and organisational learning (collective mindfulness and its five principles: preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience and deference to expertise). Using some simple examples it is shown how this framework can be used in practice, as a tool to improve the understanding, assessment, management and communication of risk. Further research and testing are, however, required to be able to use this framework in practical settings. Such research and testing are currently being conducted in the Norwegian oil and gas industry.

References

1. Althaus CE (2005) A disciplinary perspective on the epistemological status of risk. Risk Anal 25(3):567–588
2. Aven T (2000) Risk analysis — a tool for expressing and communicating uncertainty. In: Proceedings of ESREL 2000, Balkema Publishers, Edinburgh, 21–28
3. Aven T (2009) Safety is the antonym of risk for some perspectives of risk. Saf Sci 47(7):925–930
4. Aven T (2011) On some recent definitions and analysis frameworks for risk, vulnerability, and resilience. Risk Anal 31(4):515–522
5. Aven T (2012) The risk concept — historical and recent development trends. Reliab Eng Syst Saf 99:33–44
6. Aven T (2013) Practical implications of the new risk perspectives. Reliab Eng Syst Saf 115:136–145
7. Aven T (2013) Probabilities and background knowledge as a tool to reflect uncertainties in relation to intentional acts. Reliab Eng Syst Saf 119:229–234
8. Aven T (2014) Risk, surprises and black swans. Routledge, Abingdon
9. Aven T (2015) The concept of antifragility and its implications for the practice of risk analysis. Risk Anal 35(3):476–483
10. Aven T, Baraldi P, Flage R, Zio E (2014) Uncertainty in risk assessment: the representation and treatment of uncertainties by probabilistic and non-probabilistic methods. Wiley, Chichester
11. Aven T, Kristensen V (2005) Perspectives on risk: review and discussion of the basis for establishing a unified and holistic approach. Reliab Eng Syst Saf 90(1):1–14
12. Aven T, Krohn BS (2014) A new perspective on how to understand, assess and manage risk and the unforeseen. Reliab Eng Syst Saf 121:1–10
13. Aven T, Renn O (2009) On risk defined as an event where the outcome is uncertain. J Risk Res 12(1):1–11
14. Aven T, Renn O (2010) Risk management and governance: concepts guidelines and applications. Springer, Heidelberg
15. Aven T, Zio E (2011) Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliab Eng Syst Saf 96(1):64–74

16. Ambrose F, Ahern B (2008) Unconventional red teaming. In: Anticipating rare events: can acts of terror, use of weapons of mass destruction or other high profile acts be anticipated?
17. Beck U (1992) Risk society: towards a new modernity. Sage, London
18. Bergman B (2008) Conceptualistic pragmatism: a framework for Bayesian analysis? IIE Trans 41(1):86–93
19. Ben Ale JM (2002) Risk assessment practices in the Netherlands. Saf. Sci. 40:105–126
20. Bjerga T, Aven T (2015) Adaptive risk management using new risk perspectives – an example from the oil and gas industry. Reliab Eng Syst Saf 134:75–82
21. Brown GG, Cox LA Jr (2011) How probabilistic risk assessment can mislead terrorism risk analysts. Risk Anal 31(2):196–204
22. Campbell S, Currie G (2006) Against Beck: in defence of risk analysis. Philos Soc Sci 36(2):149–172
23. Chermack TJ (2011) Scenario planning in organizations: how to create, use, and assess scenarios. BK publishers, San Francisco
24. Cox LAT (2012) Confronting deep uncertainties in risk analysis. Risk Anal 32(10):1607–1629
25. de Moivre A (1711) De mensura sortis. Philos Trans 27:213–264
26. Dekker SW (2011) Drift into failure: from hunting broken components to understanding complex systems. Ashgate, Aldershot
27. Dekker SW, Nyce JM (2004) How can ergonomics influence design? Moving from research findings to future systems. Ergonomics 47(15):1624–1639
28. Deming WE (2000) The new economics for industry, government, education. MIT Press, Cambridge
29. Douglas M, Wildavsky A (1982) Risk and culture: an essay on the selection of technological and environmental dangers. University of California Press, Berkeley
30. Rosa, EA (1998) Metatheoretical foundations for post-normal risk. J Risk Res 1(1):15–44
31. Gross M (2010) Ignorance and surprise. MIT Press, London
32. Hardy CO (1923) Risk and risk bearing. University of Chicago Press, Chicago
33. Haynes J (1895) Risk as an economic factor. Quart J Econ 9(4)
34. Holling CS (ed) (1978) Adaptive environmental assessment and management. Wiley, London
35. Hollnagel E, Woods DD, Leveson N (2006) Resilience engineering: concepts and precepts. Ashgate Publishing, Aldershot
36. Hopkins A (2013) Issues in safety science. Saf Sci 67:6–14
37. Jasanoff S (1999) The songlines of risk. Environ Value 8(2):135–152
38. Kaplan S, Garrick BJ (1981) On the quantitative definition of risk. Risk Anal 1(1):11–27
39. Kaplan S, Visnepolschi S, Zlotin B, Zusman A (1999) New tools for failure and risk analysis: anticipatory failure determination (afd) and the theory of scenario structuring. Ideation International Inc., Southfield
40. Kay NM (1984) The emergent firm: knowledge ignorance and surprise. Macmillan, London
41. Khorsandi J, Aven T (2014) A risk perspective supporting organizational efforts for achieving high reliability. J Risk Res 17(7):871–884
42. Knight FH (1921) Risk, uncertainty and profit. Mifflin Company, Houghton
43. Kunreuther H, Useem M (2009) Learning from catastrophes: strategies for reaction and response. Wharton School Publishing, Upper Saddle River
44. Latour B (2005) Reassembling the social: an introduction to actor network theory. Oxford University Press, Oxford
45. Le Coze J-C (2013) Outlines of a sensitising model for industrial safety assessment. Saf Sci 51(1):187–201
46. Leveson N (2012) Engineering a safer world: systems thinking applied to safety. MIT Press, Cambridge
47. Loch CH, DeMeyer A, Pich M T (2006) Managing the unknown: a new approach to managing high uncertainty and risk in projects. Wiley, Hoboken
48. Masys AJ (2012) Black swans to grey swans: revealing the uncertainty. Disaster Prev Manag 21(3):320–335

49. Taleb NN(2007) The black swan: the impact of the highly improbable. Random House, New York
50. Taleb NN (2012) Antifragile: things that gain from disorder. Random House, New York
51. PSAN (2013) Principles for barrier management in the petroleum industry. Technical Report 29/1- 2013, Norwegian Petroleum Safety Authority (PSA)
52. Renn O (2008) Risk governance: coping with uncertainty in a complex world. Routledge, Abingdon
53. Rosa EA (2003) The social amplification of risk, chapter the logical structure of the social amplification of risk framework (sarf): metatheoretical foundation and policy implications. Cambridge University Press, Cambridge
54. Senge P, Roberts C, Smith BJ (1990) The fifth discipline fieldbook: strategies and tools for building a learning organization. Currency, New York
55. Shewhart WA (1931) Economic control of quality of manufactured product. Van Nostrand, New York
56. Shewhart WA (1939) Statistical method from the viewpoint of quality control. Dover Publications, Washington
57. Sternberg RJ (1999) Handbook of creativity. Cambridge University Press, Cambridge
58. Turner BA, Pidgeon NF (1997) Man-made disasters, 2nd edn. Butterworth-Heinemann, Amsterdam
59. UCB (2011) Final report on the investigation of the Macondo well blowout. Technical report, University of California at Berkeley, Deepwater Horizon Study Group
60. Vatn J (1998) A discussion of the acceptable risk problem. Reliab Eng Syst Saf 61:11–19
61. Weick K, Sutcliffe K (2007) Managing the unexpected: resilient performance in an age of uncertainty, 2nd edn. Jossey-Bass, San Francisco
62. Weick KE, Sutcliffe KM, Obstfeld D (1999) Organizing for high reliability: processes of collective mindfulness, Research in organizational behaviour. Elsevier, New York, pp 81–123
63. Weiss A, Wright S (2006) Dealing with the unknown: a holistic approach to marketing and competitive intelligence. Compet Intell 9(5):15–20
64. Wintle B, Lindenmayer D (2008) Adaptive risk management for certifiably sustainable forestry. For Ecol Manag 256(6):1311–1319
65. Wynne B (1992) Social theories of risk. Risk and social learning: reification to engagement, Praeger, Westport, p 275–297
66. Zio E (2007) An introduction to the basics of reliability and risk analysis. World Scientific, Singapore

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 4
Recognizing Complexity in Risk Management: The Challenge of the Improbable

Jean Pariès

L'impuissance fondamentale de l'homme industriel à maîtriser son destin apparaît de plus en plus comme le résultat paradoxal et tragique d'une volonté de maîtrise totale—que ce soit par la raison ou par la force.

Jean-Pierre Dupuy: Ordres et désordres

Abstract In the prevailing safety management paradigm, uncertainty is the enemy, and we seek to eradicate it through anticipation and predetermination. But this strategy generates "robust yet brittle" systems, unable to handle disturbances outside their envelope of designed-for contingencies. A paradigm shift is needed. We are immersed in uncertainty. We live with it, we have evolved with it as living beings, our cognitive and social skills have developed to handle the associated unpredictability. Managing uncertainty is the way we deal with the world's complexity with our limited resources. We need to better understand these abilities and augment their power in order to better engineer resilience into our systems.

Keywords Safety paradigm shift · Uncertainty management · Resilience

4.1 Introduction

Are we on the eve of a "paradigm shift"[1] in the safety domain? Do we recognize significant limitations to the ability of the prevailing paradigm to take into account

[1]In his seminal book *The Structure of Scientific Revolutions*, Thomas Kuhn challenged the then prevailing view of Science as a continuous accretion of facts and concepts. He advocated a vision in which episodes of conceptual continuity were interrupted by 'scientific revolutions', like the Copernician one, during which failures of the current theories to properly address some observed facts are recognized, and lead to a "paradigm shift".

The fundamental incapacity of Industrial Man to control his destiny increasingly appears as the paradoxical and tragic result of a desire for total control—either by reason or by force.

The copyright to third-party material quoted in this chapter is property of its respective owners and is not included in the Creative Commons license for this chapter.

J. Pariès (✉)
Dédale SAS, Paris, France
e-mail: jparies@dedale.net

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_4

observed facts? Do we have a credible alternative? Can a different social organization and risk management consensus be built around it? In this chapter, I will argue that the prevailing safety management paradigm generates "robust yet brittle" systems, unable to handle disturbances outside their envelope of designed-for contingencies, and that a paradigm shift is needed. We need to change our vision of uncertainty. We live with it, we have evolved with it as living beings, developing cognitive and social skills to handle the associated unpredictability. We need to better understand these abilities and augment their power in order to engineer resilience into our systems.

4.2 Revisiting the Concepts

4.2.1 Limitations of the Current Paradigm

There are at least three ways for a safety paradigm to reveal its limits: (i) the safety objectives announced in its name finally prove unfulfilled in large proportions; (ii) accidents deemed impossible under this paradigm do occur; (iii) this paradigm fails to reassure the civil society. Since the Fukushima Daïïchi accident, these three paths are wide open in the world of nuclear safety. The factually observed safety level does not match with the announced objectives: the observed frequency of an uncontained core meltdown in the Western technology world is about twice the target one. In spite of frequent suspicions that Tepco has been negligent, Tepco has reasonably used the available scientific knowledge, methodology and expertise to predict the magnitude of potential tsunamis off the coast of Fukushima. What we have here is not a poor implementation of an effective paradigm, but a reasonable application of a paradigm of limited effectiveness. In other domains, like aviation, or offshore oil operations, several major recent accidents (Air France 447, Qantas 32, Deepwater Horizon) have also recently illustrated this vulnerability to the 'surprises' triggered by unexpected, unthought-of events, or simply rare events. The usual response to this kind of vulnerability sounds like: "we have taken/learned the lesson: let's add this scenario to the current threat list, and we won't be fooled next time". In other words, the usual response is a commitment to do more of what is already done: anticipating threats and responses; extending the domain of predetermination.

4.2.2 The Total Predetermination Fallacy

However, what this vulnerability clearly challenges is the safety paradigm itself, i.e. the idea that safety can be based on the expectedly exhaustive anticipation of all potential threats, and the predetermination of all the expected (safe) responses. This safety paradigm dates from the 1950s, and from the efforts made in the post-war period within the framework of the development of strategic nuclear forces, aerospace industry and nuclear electricity, to better control systems reliability and contingencies. The enemy was the emergence of unexpected behavior, including technical

failures and "human error". Methods for a systematic anticipation of hazards, failures and their effects (fault trees, FMEA), and methods for a posteriori analysis of unexpected events (root cause analysis, causal tree) were developed. And while it was quickly realized that human operators were difficult to incorporate into predictive models, the attempt to treat them similarly to technical equipment by assigning them a calculable reliability coefficient (e.g. THERP, [38]), has long been in vogue and continues today (e.g. HEART, [42]). Ideas started to evolve in the 1980s when 'soft' social sciences have been introduced in safety thinking to help better understand the role of individuals, teams, organizations and cultures in accidents. But while the focus of interest slowly shifted from technical to human failures, then from front line operators' to latent, organizational failures [32], the core of the 'safety model' has remained the same. Deviations are the modern figure of risk, and are retrospectively seen as the causes of incidents and accidents, hence systematically chased in search of the modern Grail: a world where nothing goes wrong, a perfect world.

Within this paradigm, the goal is to *reduce uncertainty*, to extend a deterministic and if not, at least a probabilistic model of the world. Uncertainty is quantified and analytic rationality is applied to demonstrate that the system is fully controllable under predetermined operating strategies, tactics and procedures. This quantification is obtained from and through the concept of risk: a combination (usually a multiplication) of the probability and the magnitude of the predictable damage. This has allowed a considerable development of safety management. However risk quantification methods do more than facilitating decisions: unnoticed to decision makers, they make decisions themselves [4]. Indeed, their multiplicative nature postulates an equivalence relationship between all kinds of risks, a "distant elephant" (a remote catastrophe) weighs no more than a close mouse (a probable small discomfort). Furthermore, not all uncertainties are calculable, and there is an epistemological break between the probabilistic and non-probabilistic randomness. And risk is always an interpretation, in a multi-dimensional, social and complex interpretive field. Risk-related decisions are necessarily the result of a political process. Risk quantification methods hence do a political job: like magicians diverting the public's attention from their trick, they focus on anticipated contingencies, on calculable uncertainties, and crush the long term uncertainty into the exponential discount of the thin Gaussian distribution tail [4, 39, 40]. They contribute to the development of a social acceptability of risk coherent with a technocratic society, in which a knowledgeable elite tells to the ignorant public: "we have the scientific knowledge, we master the only valuable rationality, so we'd rather decide for you, and don't be worried, we have full control." We think, hence you are[2]… safe!

Unfortunately, as Scott Sagan nicely put it, "things that have never happened before happen all the time". And when they happen, the superior vigilance and conformity capacities developed for the anticipated events will not help: they will often do the opposite. Ironically, the capabilities needed to cope with the unexpected are those which are eroded in the continuous attempt to prepare for the expected [29]. Consequently, 'unknown unknowns', 'fundamental surprises' [22] and 'black swans'

[2] "We think, hence you are" is the title of Nicolas Bouleau's blog.

[39] occasionally brush off the Cartesian rational construction. When the 'extremely low probability' is suddenly equal to 1, when the 'defense in depth' barriers are all submerged by the same wave, the predetermination paradigm collapses and, to use the terms of [26], reveals its fallacy.

4.2.3 What Is Uncertainty?

In cognitive psychology, decision theory, and more generally in the sciences, uncertainty is defined as a state of limited knowledge, where it is impossible exactly to describe a past event, or an existing situation, or a future outcome, or to predict which one of several possible outcomes will occur. As [24] puts it:

> There are some things that you know to be true, and others that you know to be false; yet, despite this extensive knowledge that you have, there remain many things whose truth or falsity is not known to you.

Uncertainty is sometimes differentiated from ambiguity, described as 'second order uncertainty' [35], where there is uncertainty even about the definitions of uncertain states or outcomes. The difference is that this kind of uncertainty is located within the human concepts, rather than an objective fact of nature. Similarly, the reference to uncertainty in risk management has also recently witnessed a clarification of the difference between *stochastic* uncertainty and *epistemic* uncertainty [28]. Stochastic (or random) uncertainty arises from the intrinsic variability of processes, such as the size distribution of a population or the fluctuations of rain fall with time. Epistemic uncertainty arises from the incomplete/imprecise nature of available information and/or human knowledge. When uncertainty results from of a lack of *obtainable* knowledge, it can be reduced by gaining more knowledge, for example through learning, database review, research, further analysis or experimentation. But uncertainty can also result from a more *fundamental limitation* of potential knowledge. Such limitations may apply to observation, even in the 'hard sciences': in quantum mechanics, Heisenberg's Uncertainty Principle states that an observer cannot know both the position and velocity of a particle. It can also apply to the understanding process itself. The dominant scientific explanation is reductionism, which consists of decomposing phenomena, systems and matters into interacting parts, explaining properties at one level with laws which describe the interaction of component properties at a lower level. But can we 'explain' all the properties of the world (physical, biological, psychological, social…) through such a reduction process? It could be that we could *in principle*.[3]

But the least one can say is that the reductionist strategy does not have the same usefulness for all aspects of the world. Even apparently simple systems with very

[3]The French mathematician and astronomer Pierre Laplace once nicely captured this vision. His contention was that the states of a macro system are completely fixed once the laws and the initial/boundary conditions are specified at the microscopic level, whether or not we (limited) humans can actually predict these states through computation. This is one form of possible relationship between micro and macro phenomena, in which the causal dynamics at one level are entirely determined by the causal dynamics at lower levels of organization.

few components can actually be "complex systems" [23] and exhibit behaviors that cannot be predicted. One of the simplest of these systems is the double pendulum. Its behavior is called a '*deterministic chaos*': it is deterministic (the future is fully determined by initial conditions) yet unpredictable on the long term, because tiny differences in these initial conditions generate widely diverging outcomes. Lorenz [25] defined chaos as "when the present determines the future, but the approximate present does not approximately determine the future". Systems with a large number of components can also exhibit unpredictable behaviors. In *simple*, large scale systems, microscopic complexity (individual variety, or in other words, noise) disappears at larger scales because local mutual interactions annihilates individual variations (e.g. in a fluid, at the macro level, the Brownian movement of particles is "submerged" by statistical mean values, and the liquid is still). On the opposite, in large scale *complex* systems, local variety (small differences) is combined and amplified. Nearby states diverge from each other at an exponential rate, which makes these systems highly sensitive to initial conditions. Microscopic complexity hence *creates properties* at larger scales: the "noise" determines the system's large-scale behavior. Sand piles, weather, and insect colonies are examples of such chaotic behavior. This broader form of relationship between micro and macro levels, in which properties at a higher level are both dependent on, and autonomous from, underlying processes at lower levels, is covered by the notion of *emergence* [30]. Living systems but also societies, economies, ecosystems, organizations, have such a high degree of autonomy from their parts. "No atoms of my body are living, yet I am living".

Another important factor of complexity is recursion (and self-reference). The common application of recursion is defining a function or a procedure, using the function or the procedure within its own definition. Language is recursive: a dictionary defines all the words of a language, using the words of this language. Consciousness is a recursive concept, and probably one of the most complex to comprehend. Classical linear causality loses all its meaning within recursive systems.

4.2.4 Environment Ontologies: A Taxonomy of Complexity

Uncertainty appears to be the result of our cognitive limitations. Cognition, and above all, rational thinking, seems to be a process of reducing uncertainty. Knowledge, and particularly scientific knowledge, seems to be an antithesis of uncertainty. But decisions and actions are not only based on what we know, but also on what we don't know, on what we believe, on what we trust, on what we expect. Therefore it seems useful to categorize uncertainty and complexity according to the challenge posed to decision making. Several attempts have been made along these lines in the business and strategic decision domain. Courtney et al. [7] differentiate between four residual uncertainty levels (UL):

- Level 1: Quasi deterministic: only one future, with uncertainty on variants that do not change the strategy

- Level 2: A limited number of well identified possible future scenarios, each of them having a probability difficult to assess; best strategy depends on which one will actually occur
- Level 3: A continuous set of potential futures, defined by a limited number of key variables, but large intervals of uncertainty, no natural scenario; as for 2, the best strategy would change if the result was predictable
- Level 4: Total ambiguity: the future environment is impossible to forecast; no means to identify the set of possible events, even less to identify specific scenarios within this set. May be impossible to identify the relevant variables to define future.

Similarly, the 'Cynefin framework' [36] also provides a typology of contexts based on the level of complexity of the situations and problems that may be encountered. That framework intends to provide guidance about what sort of explanations, decisions or policies might apply [37]. It defines five "ontologies[4]", in other words five different types of worlds, considering their properties and level of complexity: known, knowable, complex chaos and disorder.

A third classification, based on the domain of validity of statistical methodologies, has been suggested by Nassim Taleb, and is illustrated in Fig. 4.1.

	Simple (binary) decisions A statement is "true" or "false" with some confidence interval; very true or very false does not matter. Decisions only depend on probability of events, and not their magnitude.	**Complex decisions** Both the frequency and the impact matter. There is another layer of uncertainty.
Thin tailed distributions Large exceptions occur but don't carry large consequences; "random walk" randomness; Gaussian & Poisson distributions.	Statistics does wonders. Extremely robust to black swans	Some well known problems studied in the literature (but are not numerous). Quite robust to black swans.
Fat tailed distributions Exceptions occur and carry large consequences, "random walk" or "random jump" randomness, "fractal" or Mandelbrotian distribution, unknown probabilistic structure or role of large events.	Statistical methods work surprisingly well. Quite robust to black swans.	**Black Swan domain** Do not base your decisions on statistically based claims. Alternatively, try to move your exposure type to the third-quadrant ("clipping tails"). Extreme fragility to black swans.

Fig. 4.1 The four quadrant decision-making model, adapted from [40]

[4]In artificial intelligence and information science, 'ontologies' are the structural frameworks used for organizing information and for reasoning: an ontology formally represents knowledge as a set of concepts within a domain, and the relationships between those concepts. They provide a shared semantic structure to a domain, as perceived by its actors and can serve as basis for the construction of formal reasoning or methods, and can support the design of organizations and IT tools.

4.2.5 Uncertainty and Cognitive Control

How do we deal with uncertainty? Uncertainty is the possibility for the unexpected to happen, and the unexpected is disruptive, because Human behavior is driven by "expectations". All natural cognitive systems are basically detectors of regularities, of invariant patterns or recurrent features in their environment. Human cognition developed this much further to include an ability to derive predictions from recognized past regularities which can be very abstract (principle, laws…) hence very generic. Human operators are dominantly anticipative cognitive systems. Both individual and social behaviors are driven by anticipations of the "world's" behavior. As [41] put it,

> Expectancies form the basis for virtually all deliberate actions because expectancies about how the world operates serve as implicit assumptions that guide behavioral choices.

Recent findings in neuroscience suggest that the reasoning system has evolved as an extension of the emotional system and is still interwoven with it, with emotion playing diverse and essential roles in the reasoning process. According to [8],

> Feelings are the sensors for the match or lack thereof between nature and circumstance…Feelings, along with the emotions they come from, are not a luxury. They serve as internal guides, and they help us communicate to others signals that can also guide them. And feelings are neither intangible nor elusive. Contrary to traditional scientific opinion, feelings are just as cognitive as other percepts.

Of particular interest for our discussion is Damasio's idea that the contribution of the emotional system, far from being an archaic disruptor that would degrade the performance of the reasoning process, is an essential contributor to the global cognitive performance when it comes to managing uncertainty and complexity. "Even if our reasoning strategies were perfectly tuned, it appears [from, say Tversky and Kahneman], they would not cope well with the uncertainty and complexity of personal and social problems. The fragile instruments of rationality need special assistance". This assistance is provided by emotions through "somatic markers", which "mark certain aspects of a situation, or certain outcomes of possible outcomes" below the radar of our awareness. Consequently, emotions provide instant risk assessment and selection criteria (pleasure/pain) that enable decisions and action, particularly in the presence of uncertainty. "When emotion is entirely left out of the reasoning picture, as happens in certain neurological conditions, reason turns out to be even more flawed than when emotion plays bad tricks on our decisions". So cognition does not only produce knowledge, predictability, and (provisional) certainties, it also produces an operative assessment of uncertainties, of our ignorance, of contingencies and threats that may impact our coupling to reality.

And at the same time, we simply *need* uncertainty. We need to keep some uncertainty to limit the amount of computational resources needed to act and survive in a world whose complexity is virtually infinite. Through this "bounded rationality" [34], Humans do not seek any "optimum" in either understanding or acting. They seek what is "satisficing" the achievement of their goals in the prevailing conditions.

And they constantly adjust the "sufficiency" of their behavior, hence the level of investment of their mental resources, by using heuristics rather than comprehensive analytical reasoning, adjusting trade-offs (for example, the thoroughness-efficiency trade-off [15]). In other words, they constantly manage a "cognitive trade-off" [1, 2] in order to save their mental resources. They keep as much uncertainty as possible, while remaining sufficiently effective and reliable. In order to achieve this, they "perceive" their ongoing level of control over the situation and their current and anticipated margins of manæuvre. They feel[5] when they 'control' the situation and are likely to continue doing so in the foreseeable future. Otherwise, they readjust efforts, change tactics or strategy or even higher level objectives. This ongoing perception and prediction of control is at the heart of the concept of *confidence*: it is the correct setting of confidence that allows the efficient allocation of available mental resources to the relevant issues, and thus mainly determines performance. Much of the ability to control an everyday dynamic situation is not so much in knowledge and skills than in strategies, tactics, and anticipations that allow operators to ensure that the requirements of the situation are not going to extend beyond their expertise. The talent of "superior drivers" lies in their ability to control the level of complexity and dynamics of the situation itself, so they do not have to use their (superior) skills.

In brief, our cognitive system is actually a manager of uncertainty: our cognitive system regulates the uncertainty level to handle its limitations.

4.2.6 Uncertainty and Risk Management

Far beyond that, uncertainty is also essential for evolution. Life could not exist in a totally ordered and repetitive world, because such a world would leave no room for anything new to occur. Hence randomness and uncertainty are essential for life, and a deep psychological and sociological need. We couldn't manage relationships with others if we knew everything of their feelings and intentions. Rules would not work without some dose of ambiguity and hypocrisy [3]. It is the uncertainty about the timing of our inevitable death that makes its idea bearable.

And paradoxically, uncertainty is essential for safety. The reason is that human behavior is self-represented through consciousness, hence recursive, and social behavior is more and more self-represented through the development of mass media, hence partially recursive. This opens the door to some form of downward causation, or more accurately, to a kind of retroactive causation, so to speak: A representation of the future can change the present [11]. For example, polls before a vote change the results of the vote. The paradoxical effect of the prevailing safety paradigm is to develop a collective feeling of total control, hence triggering behaviors that will reintroduce or reinforce threats. Many authors have insisted on the idea that a system

[5]This perception of control is of emotional nature, it is accompanied by a feeling of satisfaction or pleasure. Conversely, the perception of a loss of control would trigger the stress response, with the associated adrenalin. The pleasure of feeling in control is larger when the situation is inherently risky and difficult to control.

cannot be safe without some residual fear, a "reasonable dose of paranoia" as [32] put it, to keep people wary. Hence, by analogy with self-fulfilling prophecies,[6] safety can be seen as a self-defeating prophecy. The challenge is indeed symmetrical to the prophets' one: the key question is how to keep concern for risk alive when things look totally safe. As [11] put it, quoting Hans Jonas, what we need to do is to introduce "heuristics of fear" in order to "*build a vision of future such that it triggers in the present time a behavior preventing that vision from becoming real*".

This raises the issue of "whistle blowers" credibility and the issue of "weak signals". In brief a "weak signal" is a precursor of a serious problem with a low intensity or a low salience, so that it is not detected, and the potential anticipation earnings it was expectedly bearing are missed. Many efforts have been made recently to try and amplify systems' sensitivity to weak signals. However, this notion of weak signal is still fully embedded in the paradigm of anticipation, and therefore meets the same contradictions and limitations when confronted with a complex world. Hypersensitivity to initial conditions and intrinsic divergence precisely multiplies "weak signals", which are undetectable and non interpretable in real time, while crystal clear with the benefit of hindsight. Furthermore, on many occasions the signals were strong, the threat well identified, and yet the response inadequate. The problem was not the intensity of the signal but the inability of the main actors to listen to the "whistle blowers" [14]. According to [11] the main challenge may not be to *know*, but to *believe what we know*, deeply enough to act upon it and accept the usually high destabilization price, before a crisis forces us to accept.

4.3 Is There a 'Credible Alternative'?

4.3.1 Nature and Scope of Necessary Changes

As we have seen previously, the issue of uncertainty mainly boils down to the management of complexity. Basically, the contemporary prevailing safety paradigm is simplification of the world, through modeling, anticipation, predetermination. "*Obviously we can only deal with it by engaging in the vicious cycle of programming and vulnerability: more programming generating more vulnerability, which requires and legitimates increased programming. And thus a concentration of power. The self-organizing capacity of society is paralyzed, leaving a straight-growing organization*" [10]. There is a key idea in this last sentence: the idea that our societies are more and more controlled, with less and less autonomy left to individual actors,

[6] A 'self-fulfilling prophecy' is a prediction (a representation of the future) that triggers, in the present time, actions and reactions which make that future happen. Such prophecies are a central element of ancient Greece tragedies, designed to illustrate an inexorable fate. Because he was warned by an oracle that his son would eventually kill him and marry his mother, Oedipus's father abandoned his son on a mountain, hence creating the conditions for the prophecy to be fulfilled. Oedipus survived, grew as a stranger, and ignoring who they were, he killed his father and married his mother while he was trying to discover the mystery of his birth.

but also the idea that they are less and less the result of a self-organizing evolution, and more and more the outcome of a *design* process.

But *designed* systems and *evolved* systems have very different properties. One of these differences has to do with the *role of emergence*: evolved systems 'rely' on emergent structures to generate new system functionalities, while designers generally work hard to suppress emergent features. Socio-technical systems are hybrids between designed and evolved systems. They exhibit emergent properties (which by the way are not necessarily adding desirable or exploitable system functionalities). A second difference between designed and evolved systems is the *role of diversity*. For evolved systems, diversity is one of the main mechanisms that allow adaptation to changing conditions, and reversely adaptation is one of the two main processes that create diversity. For designed systems, diversity is a source of complexity, and usually seen as a source of unreliability. A third difference between designed and evolved systems is the different means used to achieve functional robustness. Edelman and Gally [12] argue that designed systems rely on redundancy, while evolved systems are characterized by a degeneracy of the structure-function mapping: biological structures typically have more than one functionality, and biological functionality is achieved by more than one structure, in very different ways. They also develop functional vicariance, which means a flexible/adaptable structure-function mapping: a structure dedicated to one functionality can be reconverted to achieve a different functionality if need be.

Hence a challenger paradigm should propose both a way to get out of the simplification-vulnerability circle, hence "*outmaneuver complexity*", and a way to organize the socio-technical systems differently, reintroducing a proper account of the self-organizing forces.

4.3.2 Suggesting New Trails

4.3.2.1 Resilience Engineering

Resilience has been defined as "the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions" [17, 18]. This notion is therefore particularly relevant to think about the management of uncertainty. There are two main ways to understand resilience. The first one is to think in terms of control, the second one is to think in terms of adaptation.

In the first perspective, socio-technical systems are regarded as (self)-controlled systems, with a combination of open and closed control loops. Hollnagel [16] proposes a list of four "cornerstones" underlying resilience, with a view that is close to the robustness of its control function with respect to disturbances: the ability to *react* (the system must "know what to do" to respond appropriately in real time, including to the unexpected); the ability to *monitor*: (the system must be able to know "what to watch" to detect potential threats, monitor its own internal state, the state of its processes, and its environment, in order to maintain the necessary regulations

to fluctuations, and to detect destabilizations that require a change of functioning mode); the ability to *anticipate*: (the system must have a sufficient phase advance, so to anticipate what will happen, predict the development of potential threats or opportunities, and more generally predict changes and their consequences, in order to maintain sufficient leeway to different time horizons); the ability to *learn*: the system must be able to expand its repertoire of responses based on its experience, but also to adapt its anticipatory strategies on the basis of the success or failure of past strategies.

In the second perspective, organizations are regarded as complex adaptive systems. The question is how to preserve or develop the adaptive capacity. In the ecological field, [13] identified four critical factors that interact across spatial and temporal scales and that seem to be required to deal with the dynamics of natural resources during periods of change and reorganization:

- Learning to live with uncertainty. The existence of unpredictable surprises should be accepted as normal. Management should not seek to systematically eradicate disturbances, but rather to deal with their effects by spreading risk through diversification of resources, operating modes and activities.
- Maintaining internal diversity (including varieties with low yields) to facilitate reorganization and renewal, investing in emergency stocks and reserves, selecting components tolerant to disturbance.
- Combining several types of knowledge for learning, which enables social actors to work together, even in an environment of uncertainty and limited information. Scientific understanding can be enriched by explorations of local and traditional knowledge.
- Create opportunities for self-organization. Self-organization connects the previous three factors. It increases the capacity for interaction between diversity and disturbances. It nourishes the learning process, and roots it in a continuous process of brewing and trial and error of institutional arrangements and knowledge, enabling them to be tested and revised.

Complex adaptive systems are generally 'robust yet fragile' [6], and caught in a 'spiral of complexity', a permanent race between complexity production and complexity needs. They are always partly undetermined and maladapted to their environment, and their adaptability results from this partial inadequacy. Dupuy [10] insists on this last issue:

> If a self-organizing system is capable of generating novelties, it is because it has the ability to adapt to random events that disturb it, to assimilate them by changing its structure. […] But for this […] a necessary condition […] is a degree of undeterminacy: it is, among other conditions, because the system is partially undetermined that it can integrate disturbances that affect it, and transform them into meaningful experiences.

All complex adaptive system strive to manage a balance between their degree of specialization and short-term optimization, and the robustness of their performance outside their adaptation envelope, that is to say, a compromise between optimality and fragility. The more a system is "fit" (optimized for a given equilibrium), the

more sensitive it will be to disturbances that may disrupt this balance. Resilience engineering has to do with the proper management of this optimality–brittleness trade-off, in other words, with the maintenance of the adaptive capacity of a system, whether it is first-rate adaptation (ability to maintain a balance) or second order (the ability to change and develop coping mechanisms to find a new balance). Hollnagel et al. [18] consequently describe the resilience of a system as being able to:

- Recognize the signs that its adaptability is falling or inadequate given the current level of uncertainty and constraints, and given future bottlenecks;
- Recognize the threat of depletion of its reserves or buffers;
- Identify when it is necessary to change priorities in the management of trade-offs, and to the adopt a higher-order logic;
- Change perspective, contrast perspectives beyond nominal system states;
- Browse the interdependencies between roles, activities, levels, objectives;
- Recognize the need to learn new ways to adapt;
- Analyze its modes of adaptation and risk assessment;
- Maintain its ability to adapt, generate and constantly regenerate its potential to adapt to a spectrum of situations as broad as possible.

4.3.2.2 Organizational Aspects

The High Reliability Organizations (HRO) community (e.g. [19–21, 33]) tried to define the features shared by organizations that seem to be "highly reliable" in their management of safety. The classical model of the pyramid organization is a homeostatic hierarchical control model, where the control centre (at the top) regulates the conditions of process coupling to the real environment (at the bottom) to reach its objectives by compensating for the variations it is subjected to. The HRO trend has shown that organizations capable of maintaining their reliability levels do not follow such a bureaucratic hierarchical structure, but are rather characterized by both a powerful centralized and strategic decision making process (i.e. consistent with the classical hierarchical model), and a powerful decentralised operational decision making process, which confers on operators at the bottom a strong empowerment, for safety issues in particular.

Along the same lines, the collibrationist[7] movement [9], considers that in risk management, the idea that societies formulate objectives based on acceptable levels of risk and seek to achieve them through a rational management is a fiction, especially because there is no social rationality to define an 'acceptable level of risk'. Therefore risk regulation is not a homeostatic process regulating a target, but the outcome of the game between many antagonist forces representing the interests of different stakeholders, through a process of 'coopetition' [5], that is to say, cooperation and competition at the same time. Rather than trying to introduce an illusory teleology, collibrationists advocate institutionalizing a game of 'tug of war' between

[7]The term co-libration comes from the English verb "to librate" meaning to give a little tap on the pan of a balance to check that it has reached equilibrium.

the opposing tensions, so that they are included in the institutions, that all parties are identified and represented, that the conflicting values involved are publicly debated, and that all interests can defend themselves, so that the final balance is not found by crushing one of them. Control is exercised by the regulatory power by changing the constraints in action in balancing mechanisms, for example by adjusting the regulatory constraints, the taxes, the access to information, etc. in order to maintain an equilibrium between powers.

Such 'polycentric structures'[8] have a number of favorable characteristics for resilience. They generate more cross-checks, a better balance of forces, a better discussion and a more open competition of ideas, more alternatives in case of failure of the ongoing policy. In long-term, this improves the management of transactions and compromises, the distribution of responsibilities, and the coordination of behaviors. In particular, as institutions are established by those familiar with local conditions, they have a better fit and a greater adaptability to these conditions, a greater legitimacy and a higher level of participation.

At the price of redundancy and of some apparent short term waste, polycentric governance is seen by its supporters as globally more efficient, especially in the long term, to manage conflicts of interest towards common resources (the 'tragedy of commons'), or between conflicting goals, especially when the characteristic time horizons are not the same between the different interests. Organizations usually have multiple and partially contradictory objectives, which can vary depending on the components (e.g. various job profiles) and also have different time-frames. Organizations attempt to balance their performance to achieve these various objectives, and thus, in a bounded world, tradeoffs are necessarily at stake. Resilience somehow measures the quality and robustness of these tradeoffs, i.e. their stability in the presence of disturbances. In this respect, another important resilience characteristic relates in particular to the ability to make "sacrificing" decisions, such as accepting the failure to reach an objective in the short term to ensure another long term objective, or 'cutting one's loss' by giving up initial ambitions to save what is essential [31]. A 'sacrificing' decision is an acceptable solution found at a higher level of the means-goal abstraction hierarchy to a conflict that could not to be solved at the initial level of that hierarchy.

4.4 Conclusion

In safety management, uncertainty is seen as the enemy, and the prevailing paradigm tends to eradicate it through anticipation of all situations and predetermination of corresponding responses. But uncertainty is everywhere, and the current safety strategy generates a vicious cycle of predetermination and vulnerability, more predetermination generating more vulnerability, which requires more predetermination, hence "robust yet brittle" systems, less and less able to handle disturbances outside

[8]The notion of 'polycentric governance' was generalized by [27], who earned the Nobel Prize in 2009, for her work on the governance of complex 'socio-ecological' systems that share common resources (such as water).

their envelope of designed for contingencies. This will inexorably lead to more 'fundamental surprise' accidents. A paradigm shift is needed. Another approach to safety is possible. Uncertainty is not necessarily bad. Actually we are immerged in uncertainty, we live with it, and we need it to deal with the world's complexity with our limited resources. We have inherited cognitive and social tools to manage it and deal with the associated unexpected variability. We need to better understand these tools and augment their efficiency in order to engineer resilience into our socio-technical systems.

References

1. Amalberti R (1996) La conduite de systèmes à risques, 2nd edn. Coll. Le Travail Humain. PUF, Paris
2. Amalberti R (2001) La maîtrise des situations dynamiques. Psychologie Française 46(2):105–117
3. Bieder C, Bourrier M (eds) (2013) Trapping Safety into Rules - How Desirable or Avoidable is Proceduralization?. Ashgate, Farnham
4. Bouleau N (2009) Mathématiques et risques financiers. Odile Jacob
5. Brandenburger AM, Doubleday BJ (1997) Co-opetition. Currency Doubleday, New York
6. Carlson JM, Doyle J (2002) Complexity and robustness. Proc Natl Acad Sci 99:2538–2545
7. Courtney H, Kirkland J, Viguerie P (1997) Strategy under uncertainty. Harv Bus Rev 75(6):66–79
8. Damasio AR (1994) Descartes' error: emotion, reason, and the human brain. Putnam Publishing, New York
9. Dunsire A (1993) Manipulating social tensions: collibration as an alternative mode of government. Technical report, Max-Planck-Institut für Gesellschaftsforschung, Köln. MPIFG discussion paper 93/7
10. Dupuy J-P (1982) Ordres et désordres: enquête sur un nouveau paradigme. Seuil, Paris
11. Dupuy J-P (2004) Pour un catastrophisme éclairé: Quand l'impossible est certain. Seuil, Paris
12. Edelman GM, Gally JA (2001) Degeneracy and complexity in biological systems. Proc Natl Acad Sci USA 98(24):13763–13768
13. Folke C (2006) Resilience: the emergence of a perspective for social-ecological systems analyses. Glob Environ Chang 16(3):253–267
14. Gilbert C (2013) Entendre aussi les signaux forts. Prévenir les crises. Armand Colin, Paris
15. Hollnagel E (1999) The ETTO principle: efficiency-thoroughness trade-off - why things that go right sometimes go wrong. Ashgate, Farnham
16. Hollnagel E (2009) The four cornerstones of resilience engineering. Resilience engineering perspectives. Volume 2: Preparation and restoration. Ashgate, Farnham, pp 117–134
17. Hollnagel E, Woods DD, Leveson N (2006) Resilience engineering: concepts and precepts. Ashgate Publishing, Aldershot
18. Hollnagel E, Pariès J, Wreathall J (2011) Resilience engineering in practice: a guidebook. Ashgate, Farnham
19. La Porte TR (1996) High reliability organizations: unlikely, demanding, and at risk. J Conting Crisis Manage 4(2):60–71
20. La Porte TR, Consolini P (1991) Working in practice but not in theory: theoretical challenges of high-reliability organizations. J Public Admin Res Theory 1:19–47
21. La Porte TR, Rochlin G (1994) A rejoinder to Perrow. J Conting Crisis Manage 2(4):221–227
22. Lanir Z (1986) Fundamental surprises. Decision Research, Eugene
23. Le Moigne J-L (1990) La modélisation des systèmes complexes. Dunod, Paris

24. Lindley DV (2006) Understanding uncertainty. Wiley-Interscience, New York
25. Lorenz EN (1963) Deterministic non-periodic flow. J Atmosph Sci 20(2):130–148
26. Mintzberg H (1996) The rise and fall of strategic planning. Free Press, New York
27. Ostrom E (2010) Beyond markets and states: polycentric governance of complex economic systems. Am Econ Rev 100(3):641–672
28. Hoffman FO, Hammonds JS (1994) Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 14(5):707–712
29. Pariès J (2011) Lessons from the Hudson. Resilience engineering in practice: a guidebook. Ashgate, Surrey
30. Pariès J (2006) Complexity, emergence, resilience. Resilience engineering: concepts and precepts. Ashgate, Aldershot
31. Pariès J (2012) Resilience in aviation: the challenge of the unexpected. Presentation at IAEA Technical Meeting on Managing the unexpected from the perspective of the interaction between individuals, technology and organization
32. Reason J (1997) Managing the risks of organizational accidents. Ashgate, Farnham
33. Rochlin GI, La Porte TR, Roberts H (1987) The self-designing high-reliability organization: Aircraft carrier flight operations at sea. Naval War College Rev 40(4):76–90
34. Simon HA (1982) Models of bounded rationality: behavioral economics and business organization. MIT Press, Cambridge
35. Smithson M (1989) Ignorance and uncertainty: emerging paradigms. Springer, New York
36. Snowden DJ (2005) Multi-ontology sense-making, a new simplicity in decision making. Inf Prim Care 13:45–53
37. Snowden DJ, Boone ME (2007) A leader's framework for decision making. Harv Bus Rev 85(11):68–76
38. Swain AD (1972) Design techniques for improving human performance in production. Industrial and Commercial Techniques Ltd, London
39. Taleb NN (2007) The black swan: the impact of the highly improbable. Random House, New York
40. Taleb NN (2008) The fourth quadrant: a map of the limits of statistics. Edge
41. Weick KE, Sutcliffe KM (2001) Managing the unexpected: assuring high performance in an age of uncertainty. Jossey-Bass
42. Williams JC (1985) HEART - a proposed method for achieving high reliability in process operation by means of human factors engineering technology. In *Proceedings of a symposium on the achievement of reliability in operating plant, safety and reliability society (SaRS)*, Birmingham

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 5
Practices in the Danger Culture of Late Industrial Society

Arie Rip

Abstract The chapter replaces the question of risk control by one about how we handle danger in our societies and realize a measure of safety. Ongoing practices in a framework of 'danger cultures' are the key. The case of environmental and health inspection and the intersecting 'social worlds' involved, are used as a case to indicate important features.

Keywords Danger culture · Practices · Repertoires · Regulation · Good firms

5.1 Introduction

The world is full of dangers, actual and potential. People and societies can be fatalistic about them, but also attempt to reduce the dangers, for example by building dikes (and maintaining them) as a defense against floods—an example that comes easily for somebody from the Netherlands. Closer to home is how one's house is supposed to keep one safe from the dangers outside.[1]

Professional specialities have evolved, up to safety engineering. In modernity, and particularly with the rise of the regulatory state since the late 19th century, an overall idea of protection against dangers has evolved, up to ambitions of controlling risks.[2] This can be traced in some detail, including the rise of probabilistic risk analysis [16].

[1] Above the entrance to a big house near the Vliet in the Netherlands is the inscription "In de wereld is veel gevaar" (There's lots of danger in the world), indicating that you will leave the dangers behind when you enter the house. The house (like the cave in prehistoric times) protects you against the dangers, the wild animals roaming outside.

[2] Cf. [6], who tells the story of attempts to develop probability, based on data from the past. He comments, in his concluding chapter: "The past seldom obliges by revealing to us when wildness will break out in the future". (p. 334). Continuing the observation in the preceding footnote, one could say that the future is full of wild animals, and that risk control is an attempt to tame the future.

A. Rip (✉)
University of Twente, Enschede, The Netherlands
e-mail: a.rip@utwente.nl

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_5

The present book addresses the illusion of risk control, at least in so far as the attempt at control is aimed at reducing uncertainty. But the issues are just as much about safety and about danger, and protection from danger—up to illusions of protection. This move has been visible in the literature on safety and safety culture (for example [4, 11]) and has led to critical reflection [7].

It is important to broaden the picture in this way, not just because risk is a quite specific concept (cf. [16]), but also because there are problems with the notion of control. The illusion of risk control is the illusion of a priori risk control, as if there were no further developments and contingencies, and uncertainties could be made manageable (so that we could feel safe). As I will argue, risk/danger control is possible, but not a priori and independently of local contexts. But the illusion of risk control cannot be given up, because governments and managers have to prepare for eventualities, and do so by setting rules and regulations in the hope that these will have some effect.

Coming into these questions from the other side, as it were, from the side of danger and safety, one can consider actual practices to reduce danger and ensure some safety in an uncertain world, not always visible to the outside, but important to recognize for what is happening in them. One can be concerned by the risks of late industrial societies (cf. also [5] on the risk society), as well as in less developed countries, and rightly so. But with the uncertainties and increasing complexity of our societies, and the human propensity to focus on own immediate interests and not take up wider responsibilities, it is actually remarkable that there are not more accidents and damages of industrial and transport activities. There must be 'repair work' going on, as in the 'near accidents' (as with trains and planes, but also in road traffic).

Inspired by some work in Science and Technology Studies, I develop a perspective on ongoing practices of local reduction of unsafety and danger, which are all part of the 'danger culture' of our industrial society—which includes regulation [17]. This perspective fits perfectly in the critical reflection on proceduralization of safety [7], but adds a further element, the intersection of organisations and "social worlds" and the role of professionals [18].

5.2 The Danger Culture of Industrial Society

The notion of 'danger culture' was introduced by [20] to characterize the world of miners and mountaineers (and one could add, divers and other hazardous occupations and pastimes), where technical precautions and rules (including rules of solidarity) evolve and are enforced by the actors themselves.[3] In doing so, the members of these

[3]It is useful to quote Turner at some length, to indicate his notion of danger culture as applying to a subculture rather than a whole society, which is my point: "Both the isolation and the danger which characterize such occupations promote cohesion and self-sufficiency, a sharing of danger and a suspicion of outsiders who do not share it. And all of these characteristics contribute towards the development of an occupational subculture of danger [...] which is similar in some ways to

worlds can push into the background the essential uncertainty with which they live. Thus, there is a dual survival strategy: reducing unsafety and dangers in daily life, in order to be able to continue to live in a world that is essentially dangerous.

At the macro-level of our industrial society, one sees a version of such a dual survival strategy. There are regulations of various kinds, based on expert advice (and now also including a precautionary component). The general trust in the effectiveness of regulations allows people to background the overall uncertainties of living in our late-industrial society. There might be a tinge of fatalism in this trust, linked to the delegation of responsibility to authorities and official regulations, which can turn into cynicism.

Ulrich Beck's diagnosis of the 'risk society' overlaps with my diagnosis of the danger culture of late-industrial society, but emphasizes the big risks and the organized irresponsibility that appears to be our way of handling them. He focuses on high-tech risks (nuclear, chemicals). But many actual dangers and their deadly effects derive from neglect, in the small and in the large.[4] Modest practices of vigilance, care, and what is called 'repair work' in sociology,[5] will be more important to maintain safety than regulations (which can never capture the local specificities and contingencies that determine actual outcomes).

I tend to argue against too much trust in regulation, and instead, put local practices up front when considering safety. But however important the focus on the local is, it cannot be sufficient, because some risks derive from the extra-local and from the 'not yet' [1].

Danger culture takes different forms and plays out differently at different levels. The subtitle of the original workshop theme paper asked: what does it take to live with uncertainty? I discuss what it takes to live with danger, from the known but uncertain dangers to the unknown unknowns. I will do so by briefly elaborating on the danger culture of industrial society, then present my perspective on "intersecting social worlds" [18] which is one essential element in how we live with danger.

---

the disaster subculture developed amongst those whose homes are frequently exposed to natural hazards [...]. The subculture provides shared perspectives which enhance the group's control over their work situation, which maximise autonomy and minimise dependence upon outsiders. Fear is often, though not always, denied within the group as a way of making the internal environment more predictable, and group norms are very important both in controlling and testing new members and in (Footnote 3 continued)
encouraging them to behave predictably in the face of danger. [...] Safety may be taken seriously at one level by regarding all possible hazards as extremely dangerous (though this may not include all conditions seen as hazardous by outsiders), but these occupational groups differentiate in practice between different kinds of hazard. Miners, for example, give most instruction to newcomers about everyday dangers such as falls and slips, less about dangers of an intermediary kind such as misfired explosions, and in the face of the possibility of major hazards such as fires and large cave-ins they are more stoic and fatalistic." [20, p. 67].

[4]The April 2013 collapse of the Rana Plaza garment factory complex in Bangladesh, causing the death of more than a thousand workers, is a dramatic example.

[5]Grote [10] makes the same point when arguing that the socio-technical model of safety culture emphasizes the importance of "supporting local actors in controlling variances at their sources" (p. 645).

Even while cultural rules and repertoires tend to be reproduced, practices in particular danger cultures (and more generally) evolve. This is visible at the meso-level, in the communities of miners and mountaineers. Explicit rules may emerge, covering a variety of situations, and these may be formalized and become part of the official rules of the organization and/or the sector. But then, at the micro-level, in specific settings, there are ways of doing that keep the practice a going concern even when this goes against the rules. The latter may avoid danger in concrete cases (cf. 'repair work'), but also undermine safety in the longer term (cf. the example of [21] of alarms in nuclear plants that are too sensitive, so people learn to ignore the alarm signals and/or dismantle the devices – until such time as there is real cause for alarm but it is not noticed or followed-up).

At the macro-level, danger culture includes prescriptions and requirements (for safe processes and products), the regulation of chemicals, safety rules in traffic. Thus, a network of rules (like taboos) that allows us to continue with what is essentially dangerous business: living in industrial society [8].

Safety, in the sense of "avoidability of unacceptable risk" (to paraphrase the original workshop theme paper), refers to practices and to responses to events as they occur as much as to anticipatory knowledge feeding into regulation. I will use a concrete case to discuss this further: the explosion of a fireworks factory in the city of Enschede on May 13, 2000, devastating the immediate surroundings and killing 18 people.[6] It is not completely clear what the sequence of events and the causalities were, and where blame might be located. The case does allow identification of three components of safety in practice.

There are technical aspects, in this case that the concrete covers of the bunkers in which the fireworks were stored were too heavy, so that more pressure built up inside the bunkers, and the eventual explosion created more damage (through the shock wave, and through blocks of concrete flying through the air) than necessary. Afterwards, military specialists in storage of explosives pointed this out, referring to their own rules. This is just one technical aspect. In general, there are issues of reliability, strength of constructions, measuring health and environmental effects of chemicals.

Then, there are contingencies on location, in particular the absence of 'repair work' when there was a small fire in the yard of the factory (intentionally or unintentionally), and the few people around on that Saturday morning/early afternoon let it continue. This has led to attempts to find fault, and blame people. Conversely, the unavoidable "uncertainty about sequences of events" should be an occasion to push for watchfulness and repair work of the actors, which should be part of the culture (as it is with miners and mountaineers).

Thirdly, there had been a certain laxness of the city's inspectors in applying official rules to what the factory was doing, situated in the midst of a built-up area. After the fact of the explosion, this was criticized, but it is part of a general aspect of control of risks. To assure safety, it is not enough to focus on compliance to stipulated regulation,

[6]This event has been referred to in the literature [12, p. 271]. I add my own involvement in that I lived uncomfortably close by, but experienced only extensive material damage to my house.

and rules more generally, and it may be counterproductive in terms of actual safety. The actual behaviour of firms is the important thing, and actual interactions between safety inspectors and firms have focused on this. If the firms tend to be "good" firms, then, if they happen to not follow a rule, or an accident occurs, this need not be sanctioned because they generally try to do a good job. Clearly, this creates a "grey zone" of interactions and negotiations, and things may go wrong.

It is this last point that I want to develop further, using workplace and environmental safety as the concrete domain. While my discussion builds on some of the specificities of that domain, like the difficulty of monitoring and the problems of sanctions, the point about intersecting social worlds and the role of C- and R-repertoires (see below) is quite general and applies to all sorts of professionals and proto-professionals. In other words (and in line with some recent literature like [10] and [7]), important aspects of risk and safety are about general social and behavioural dynamics, rather than specifically about risk and safety. The recent interest in 'safety culture', however important in its own right, is still limited because of its strong links to management questions of how to create and maintain it [14], in a further step leading to blaming organizations for not paying sufficient attention to their safety culture, rather than inquiring into actual safety. In my discussion below, I draw attention to interactions across organizations (that is why I refer to 'social worlds') their nature and quality in terms of assuring some safety. 'Danger culture' is an even broader concept. In this chapter I limit myself to how it plays out in the worlds of professionals.

5.3 Negotiations in Intersecting Social Worlds, Rather than Implementation of Regulation[7]

Let me start with a puzzle. Many industrial firms comply with environmental regulations. However, it is is not obvious that they will do so. If the profit motive is what drives them (and managers and CEOs tend to refer to the need to make a profit, if only to close difficult discussions about what the firm should do), why would they ever comply to environmental regulations, e.g. about handling wastes? Violations (infringements) of the rules are unlikely to be discovered (in the 1980s, only 25 % of environmental violations were detected, and when found out, the sanctions are light and/or can be postponed and reduced by protests and court cases. A simple cost-benefit calculation would drive firms to not, or only minimally, comply with regulations. Indeed, there are so-called 'cowboy firms' which do not comply at all, and enjoy the benefits as long as they can. If they are caught, and face strong action, they might just end their business (and continue their operations elsewhere, under

[7]This section draws heavily on an unpublished paper, which is now in the public domain as part of a collection of papers by Arie Rip on the occasion of a conference in honour of his being retired Rip2011. The text and figures draw on this paper, with only minor modifications.

another name). But the majority are 'good firms', who want to avoid scandals, and pride themselves on maintaining good relations with environmental inspectors.[8]

What happens is that the good firms, or at least parts of them, are part of a new social world, together with the environmental inspectors out on duty in the field. It is an enforcement and compliance world, and a key component is how a productive C-repertoire has emerged, with linkages to the R-repertoires of the firms (profit motive) and the inspectorate (enforcement of regulation).[9] Here, I am using concepts inspired by science and technology studies on the difference between the discourse within the protected spaces of science, and the discourse in presenting and justifying the work of science to the outside world [9, 18]. In the contingent (C) repertoire, informal interactions and negotiations can be conducted in local practices, and are accepted. The rational (R) repertoire has formal rules and justifications, for outside use, but can also be referred to within the C repertoire as a sort of boundary condition. Professional practices of lawyers, scientists and medical professionals can be analysed in those terms.[10]

The C-repertoire allows inspectors to forget about enforcing the formal rules and focus on avoiding environmental pollution in practice, in exchange for "good" behaviour of the firms they interact with. "Better a dirty conscience than a dirty world" is their motto.[11] That it is a social world is clear from the occurrence of inclusion and exclusion moves. Firms which go against the informal rules of good behaviour are labeled as deviants ("cowboy firms"), and are treated harshly. "Good firms" on the other hand can have their occasional waste problem, but it is then treated as an unfortunate accident. At the side of the inspectorate, there are also inclusion/exclusion pressures. They have to avoid strict adherence to the rules, or will be seen as fanatics (also by their colleagues) and disavowed.[12]

---

[8]The distinction between 'good firms' and 'cowboy firms' returns again and again, and is not limited to firms. NRC Handelsblad, 29 April 2006, presented commercial stem cell therapy ("trade in hope") under the heading: 'Stem cell cowboys conquer the world, now also in Rotterdam'. On the front page, it referred to "dubious treatments", and used as heading: "Specialists call for a stop of 'stem cell pirates'. The "good" therapists were trying to exclude the "cowboys".

[9]Here, and subsequently, I focus on interactions between firms and inspectors, but the same type of analysis is applicable within an organisation. See for example the empirically rich report on incident reporting in a nuclear research centre [19].

[10]See for example [2] on the world of court rooms.

[11]It was their motto, at the time. Changes in governance (partly inspired by new public management) have forced inspectors to become more distantiated, with fewer interactions in the field. A striking example is the Dutch Occupational Health and Safety Inspectorate, which is now limited to advise on health and safety activities in general, rather than interact with firms.

[12]Here and in the following, I use data and quotes drawn from a Special Issue of Policy Studies Journal (1982) to illustrate these points, drawing on a draft paper for a summer workshop at IIASA, July 1983. There is a contingent repertoire (C) which talks of bargaining, of being pragmatic, "the 'twilight zone' process of seeking voluntary compliance and negotiating stipulations" (59) [Numbers between parentheses refer to pages of Policy Studies Journal.] There is also a rational repertoire (R) where enforcement is seen as the execution of the rules, "command" instead of "bargaining" (139), the "strict liability" that does not take into account the intentions of the actors (160) or the possibility of "accidents" that decrease culpability. There are two important points to note about the relation between the two repertoires. (1) Inspectors and other enforcement agents do

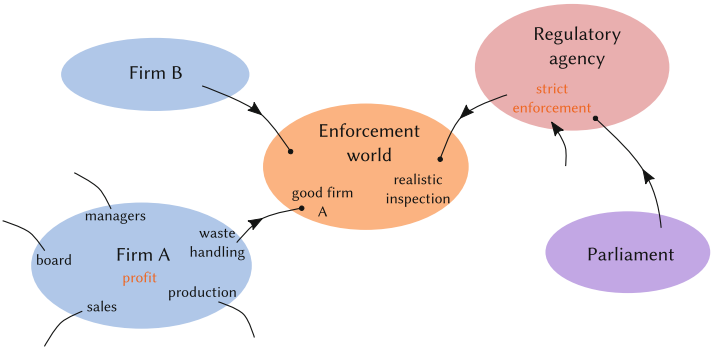

Fig. 5.1 Interlocking social worlds

Social worlds are characterized by dual (C + R) repertoires, in general, but this also drives newly emerging "bridging" social worlds, like the enforcement and compliance world which bridges firms (certain departments and individuals in firms) and inspectors. In the C-repertoire, negotiation and interpretation are central. The inspector checks whether the firm remains within certain levels of compliance, and interprets violations as accidents for which the firm can be excused. The inspectors can refer to the official regulations and sanctions—that is their R-repertoire—but only as a last resort. If they were to apply the regulations literally, that might lead to system-wide protests and refusal (one could see this as a form of civil disobedience), and thus be improductive.[13] Firms have their R-repertoire of profit maximisation for their shareholders, and guaranteeing continuity of the firm. Again, strict application of this R-repertoire may be counterproductive. Compare how in the 1960 and 1970s, firms like the ITT conglomerate were disavowed as focusing on profit only, and suffered from the allegation.

These social worlds, and social worlds in general, depend on the productive duality of their C- and R-repertoires, but also on their external links, through their R-repertoire or otherwise, as illustrated in Fig. 5.1. Environmental and safety staff in firms create a link with other departments in the firm and with overall management. Inspectors out in the field return to their office in the government ministry, and have to justify their actions there.

The enforcement world functions between two extremes ('poles'). One extreme occurs when the links to other social worlds are completely backgrounded, and interactions within the world are the focus. The inspectors as well as the staff of the firm "go native", their allegiance is to their shared world. This is also a way to operationalize trust. In the other extreme, the inspectors and the staff of the firm

not see it as their task to enforce the law. Instead, their goal is to protect the waste treatment system from harm, to solve effluent problems (158), to contribute to an adequate solution to the pollution problem faced in a given case while minimizing enforcement costs (139). (2) The enforcement of the law is a resource in the enforcement process, not an end in itself (163).

[13]This can be linked to the notion of 'functional forgiveness' [13].

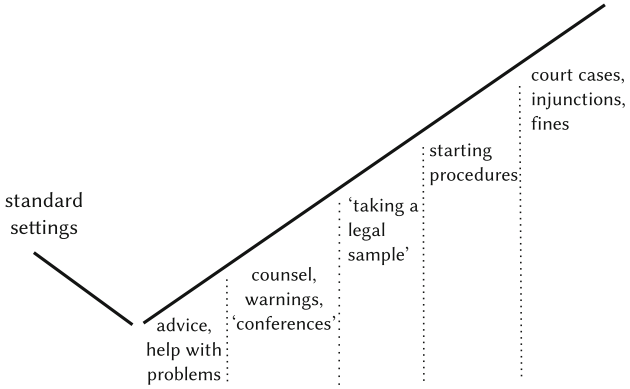

Fig. 5.2 Graded persuasion in the enforcement process

are spokespersons for their respective worlds, and they interact strategically. Since they are bound together (mutual dependency), one can speak of a strategic game and see their actions as moves in a strategic game. This is actually how the actors will interpret actions of others – seeing them as players in a game rather than members of the same social world.

In practice, mixed or compromise arrangements occur, and shift over time between the two poles. And negotiations occur, building on how inspectors construct a gradient of force to keep the firms in line. In Fig. 5.2, I visualize the gradient at the right hand side, showing the practices and measures as used at the time, going from soft to hard. The line at the left hand side about standard setting is just a reminder that the gradient emerges because of standard setting; it is not part of the graph. This visualization has the same form as the so-called enforcement pyramid, proposed by [3] to operationalize how enforcement agents can start with a persuasive enforcement style and escalate punishments only when a business consistently refuses to cooperate. It has been noted that the enforcement pyramid reflects how inspectors are inclined to behave in practice. When shown to them, inspectors perceive it as common sense, and react saying "we already do that" [15]. What I am adding is that the gradient of force only works when the firms need to remain members of the shared world, i.e. be good firms. Clearly, this is a grey zone, and one which can be productive exactly because it escapes proceduralization [7].

5.4 Conclusion

I have analysed such grey zones as a general phenomenon, especially in professional practices, up to the practice of doing science in the lab, and used analysis in terms of contingent and rational repertoires to show the structure of the situation. The point is that grey zones are not just grey and muddled, but they allow realisation

of productive outcomes in spite of conflicting, or just different, objectives, here in terms of safety and reduction/avoidance of danger. But they can also lead to neglect of repair work and create problems and accidents. That is where the R-repertoire comes in as a constraint. Not because the rules and norms in the R-repertoire are good by definition, but because they create an external reference point to avoid the C-repertoire interactions shifting towards 'anything goes'. In that sense, regulation is not about specifying what is safe and to be implemented, but like a 'regulative ideal' in the sense of philosopher Immanuel Kant in his *Kritik der reinen Vernunft*.

I am willing to argue that such grey zones are necessary to have late-industrial societies survive without too many unsafe situations and accidents, but recognize that they constitute a second-order risk, of becoming shady dealings that serve the interests of the immediately involved actors, but not the management of safety. Our overall danger culture of industrial society, as it has emerged and solidified is not of much help here. Transparency, enforced from above, may be counterproductive, as it creates countermoves, as when the USA created the "Sunshine" Act requiring all government documents to be public: there was a shift away from documents to oral interactions. More important for mitigating the second-order risk are social dynamics, in particular whistle blowing (from the inside) and critical journalism and other external engagement. When accidents or near accidents, and the resulting investigations, shine light on partly hidden practices and reveal the compromises made, this could be taken as a *prima facie* argument for transparency.[14] But transparency may reduce the scope of repair work, which is essential for actual safety. In addition, disclosure has its own social dynamics, which might constitute a third-order risk, of creating witch-hunts of practices that were actually relatively productive.

Whatever the new paradigm about risk control will be, it has to recognize the actual practices that assure (but sometimes undermine) safety, to some extent. And to understand how these play out at different levels: local practices, meso-level practices of communities/worlds, and macro-level responsibilities for regulation, as well as cultural aspects. The new paradigm, whatever else it does, should reinforce the good elements in the practices of the danger culture of late-industrial societies, based on an understanding of the patterns and dynamics at play.

References

1. Adam B, Groves C (2007) Future matters: action, knowledge, ethics. Brill, Leiden
2. Atkinson JM, Drew P (1979) Order in court: the organisation of verbal interaction in judicial settings. Macmillan, London
3. Ayres I, Braithwaite J (1992) Responsive regulation: transcending the deregulation debate. Oxford University Press, New York
4. Baram M, Schoebel M (2007) Editorial: safety culture and behavioral change at the workplace. Saf Sci 45(6):631–636

[14] I am indebted to Eric Marsden for drawing my attention to the aspect of transparency, and offering considerations like this one.

5. Beck U (1992) Risk society: towards a new modernity. Sage, London
6. Bernstein P (1996) Against the gods: the remarkable story of risk. Wiley, New York
7. Bieder C, Bourrier M (eds) (2013) Trapping safety into rules - how desirable or avoidable is proceduralization?. Ashgate, Farnham
8. Douglas M (1966) Purity and danger: An analysis of concepts of pollution and taboo. Routledge, London
9. Gilbert GN, Mulkay M (1984) Opening pandora's box., A sociological analysis of scientists' discourseCambridge University Press, Cambridge
10. Grote G (2007) Understanding and assessing safety culture through the lens of organizational management of uncertainty. Saf Sci 45(6):637–652
11. Hale AR, Swuste P (1998) Safety rules: procedural freedom or action constraint? Saf Sci 29(3):163–177
12. Jovanović AS, Renn O (2013) Search for the 'European way' of taming the risks of new technologies: the EU research project iNTeg-Risk. J Risk Res 16:271–274
13. La Porte TR. Trapping Safety into Rules - How Desirable or Avoidable is Proceduralization?, chapter Postscript: Reflections on Procedures, Trial and Error, and Functional Forgiveness, pages 257-272. Ashgate, 2013
14. LeCoze J-C, Wiig S (2013) Beyond procedures: can 'safety culture' be regulated? Trapping safety into rules - how desirable or avoidable is proceduralization?. Ashgate, Farnham, pp 191–204
15. Mascini P (2013) Why was the enforcement pyramid so influential? and what price was paid? Regul Gov 7(1):48–60
16. Rip A (1986) The mutual dependence of risk research and political context. Sci Technol Stud 4:3–15
17. Rip A (1991) The danger culture of industrial society. Communicating risks to the public: international perspectives. Kluwer Academic, Dordrecht, pp 345–365
18. Rip A (2011) Interlocking socio-technical worlds. A booklet with seven collected articles by Arie Rip, offering a cross-section of his work, at the occasion of the Symposium Futures of Science and Technology in Society at the University of Twente
19. Rossignol N (2015) Practices of incident reporting in a nuclear research center: a question of solidarity. Saf Sci 80:170–177
20. Turner BA (1981) Organizational responses to hazard. In: Kunreuther H (ed), Risk. A Seminar Series, International Institute for Applied Systems Analysis, pp 49-86
21. Wynne B (1988) Unruly technology: practical rules, impractical discourses and public understanding. Soc Stud Sci 18(1):147–167

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 6
Judicial Review of Uncertain Risks in Scientific Research

Eric E. Johnson

Abstract It is difficult to neutrally evaluate the risks posed by large-scale leading-edge science experiments. Traditional risk assessment is problematic in this context for multiple reasons. Also, such experiments can be insulated from challenge by manipulating how questions of risk are framed. Yet courts can and must evaluate these risks. In this chapter, I suggest modes of qualitative reasoning to facilitate such evaluation.

Keywords Risk · Risk scenarios · Law · Courts · Science · Scientific research · Particle physics · Black holes · LHC

6.1 Introduction

In the world of engineering, uncertainty is anathema. Whether designing a nuclear power plant or planning a system of levees, engineers do not like surprises. Experimental scientific research, however, thrives on uncertainty. Reaching beyond the current state of knowledge is the whole point of the enterprise. Whether probing the planets or trying to create new particles, scientists would love to see a surprise around every corner. Uncertainty goes hand in hand with aspirations to explore and discover.

This special role for uncertainty in scientific research makes questions of catastrophic risk from science experiments particularly interesting. What is more,

[1] See Hawai'i County Green Party v. Clinton, 980 F. Supp. 1160, 1168-69 (D. Haw. 1997); Karl Grossman, The Risk of Cassini Probe Plutonium, The Christian Science Monitor (Oct. 10, 1997) http://www.csmonitor.com/1997/1010/101097.opin.opin.1.html; CNN, Cassini roars into space, CNN (Oct. 15, 1997) http://www.cnn.com/TECH/9710/15/cassini.launch/; Najmedin Meshkati, Probability, Plutonium Don't Mix (op-ed), Los Angeles Times (Oct. 10, 2013) http://articles.latimes.com/print/1997/oct/10/local/me-41149.

The copyright to third-party material quoted in this chapter is property of its respective owners and is not included in the Creative Commons license for this chapter.

E.E. Johnson (✉)
University of North Dakota, Grand Forks, USA
e-mail: eric.e.johnson@law.und.edu

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_6

science experiments have raised the specter of catastrophe scenarios noteworthy for their exoticness and magnitude. Critics have alleged that a plutonium-laden interplanetary spacecraft may pose the risk of causing millions of cancers by disintegrating in the atmosphere.[1] A microbiological laboratory could plausibly cause a pandemic with an accidental release of an exotic pathogen [1]. And experimenting at the edge of knowledge of fundamental physics could, in the view of some, risk destroying the entire planet [11].

In this chapter, I will look at uncertainty in science-experiment risk from a legal perspective. Suppose someone brings an action in court to halt an experimental endeavor based on the allegation that it poses an undue risk to the public. Courts generally have the power to order a halt to activity that puts human life at risk.[2] But how do courts decide that there is a substantial chance of human casualties? It's not an easy task under the best of circumstances. And when the context is an activity intentionally designed to reach beyond the current state of human knowledge, it is especially tricky.

Courts have often side-stepped thorny issues of uncertainty, embracing instead the traditional view of risk.[3] In the traditional view, risk can be understood in terms of probabilities and applied in the form of cost-benefit balancing [22]. Traditional risk evaluation, then, depends heavily on experts and quantitative analysis.[4] This mode of thinking corresponds to classical theories of decision-making, in which information is available, problems are known, and optimal solutions can be found through the application of perfect reason [10, p. 123]. This traditional paradigm is broadly problematic because it ignores the existence of human limitations in understanding. In other words, it ignores the central role of uncertainty in risk.

---

[2]In common-law jurisdictions, this may be accomplished through the court's inherent equitable power. See, e.g., Harris Stanley Coal & Land Co. v. Chesapeake & Ohio Railway Co., 154 F.2d. 450 (6th Cir. 1946) (reversing district court's refusal to enjoin coal mining operations that posed a risk of subsidence to railroad tracks, because "[a] court of equity will not gamble with human life, at whatever odds, and for loss of life there is no remedy that in an equitable sense is adequate."); Shimp v. New Jersey Bell Telephone Co., 368 A.2d 408 (N.J. Super. Ch. 1976) (recognizing a "common-law right to a safe working environment" and finding that "such work area was unsafe due to a preventable hazard which the court had power to enjoin."). Courts in civil-law jurisdictions have similar powers conferred through broad provisions of the civil law.

[3]To say that courts often embrace a very traditional view of risk and side-step uncertainty is not to say they do so always. Particularly in the context of reviewing administrative regulations, courts have sometimes embraced the precautionary principle, the essence of which is to err on the side of caution in conditions of uncertainty. See, e.g., Alberto Alemanno, The Shaping of the Precautionary Principle by European Courts—From Scientific Uncertainty to Legal Certainty (Bocconi Legal Studies Research Paper No. 1007404, Aug. 16, 2007) (discussing the role played by European Community courts in shaping the precautionary principle); Indus. Union Dept., AFL-CIO v. Am. Petroleum Inst., 448 U.S. 607, 656 (U.S. 1980) ("the Agency is free to use conservative assumptions in interpreting the data with respect to carcinogens, risking error on the side of overprotection rather than underprotection").

[4][22, p. 123]. See also Chap. 4.

In many cases coming before the courts—such as garden-variety products-liability actions—the classical approach might be relatively unobjectionable. Even though uncertainty exists in all situations of risk, perhaps it is often efficient for courts to disregard uncertainty and proceed as if risk is necessarily tamable by number-crunching experts.

But when it comes to science-experiment catastrophe risk, the role of uncertainty cannot be ignored. The traditional paradigm for thinking about risk is especially problematic in this context for at least three reasons. First, reliance on experts is particularly troublesome since the only persons who can supply expert opinions on the safety of leading-edge science experiments are often the same people who are involved in the experiment, either directly or indirectly. Second, working in the certainty mode of risk assessment may require the application of scientific knowledge that the experiment itself is designed to discover, creating a kind of catch-22. Third, where the computed probability of science-experiment disaster is small, it is overwhelmed by uncertainties in the assessment. That is, the chance the probability assessment is wrong—for instance, because of erroneous assumptions, flawed calculations, or observational errors—is many times greater than the computed probability itself.

Beyond these practical and conceptual concerns, there is also a problem of rhetoric. Traditional risk assessment can be easily frustrated and manipulated by science-experiment proponents. In leading-edge science, the close-knit community of researchers championing the experiment may have an effective monopoly on expertise and knowledge in the area, and this can allow those scientists to control the framing of risk questions. This framing can be used to frustrate meaningful engagement on safety issues in terms of traditional risk analysis. And selective assertions of uncertainty can steer the debate toward a favorable outcome for experiment supporters.

I will explore these concepts primarily through one fascinating and extreme example—CERN's Large Hadron Collider (LHC), which critics say could create a planet-destroying black hole.[5] After pointing out the various troubling features of science-experiment risk, I propose a way for the courts to deal meaningfully with legal challenges to such experiments despite the many difficulties.

6.2 Background on the LHC/Black-Hole Disaster Scenario

In this part, I will provide some background on the LHC, the laboratory organization that operates it, and the black-hole risk the experiment is said to pose.

[5]The instant paper follows on a previous article I wrote that discusses some aspects of the black-hole issue in additional detail and analyzes some other issues, such as jurisdiction, evidence, injunction law, and jurisprudence [11].

6.2.1 CERN and the LHC

Located near Geneva, CERN is the world's pre-eminent laboratory for particle physics, which studies the most fundamental aspects of matter. An intergovernmental organization comprising 21 member states, CERN is a mammoth institution.[6] Its 2014 budget was 1.1 billion CHF (€1.0 billion).[7] And a recent count indicated that 10,357 experimental physicists were involved in CERN in some capacity.[8]

CERN's current program is centered around the LHC, a superconducting synchrotron particle accelerator, along with its various component experiments.[9] Launched in 2008, the LHC is, according to CERN, "the largest machine in the world".[10] The beam tunnel is 27 km around.[11] The magnets used to accelerate the particles are cooled with about 90 tons of superfluid helium. And as a whole, the apparatus consumes enough power for a medium-sized city.[12]

The aim of the LHC is to advance understanding of the fundamental particles and forces that make up the physical universe.[13] Indeed, the LHC already achieved a triumphant discovery by finding the celebrated Higgs boson—a fundamental particle that is understood to impart mass to matter.[14]

The LHC was designed to collide protons at an energy of 14 trillion electron volts (TeV), reaching a new energy range that scientists have dubbed the "terascale".[15] This collision energy will be 14 times the energy of any previous experiment.[16] As of the spring of 2015, the LHC was just beginning to get close to achieving this energy,

---

[6] See CERN, CERN brochure 1 (2010), http://cds.cern.ch/record/1278456/files/CERN-Brochure-2010-005-Eng.pdf.

[7] See CERN General Information 2014, http://cds.cern.ch/record/1745540/files/CERN-Brochure-2014-006-Eng.pdf.

[8] See CERN General Information 2012, http://cds.cern.ch/record/1453354/files/CERN-Brochure-2012-003-Eng.pdf.

[9] See CERN, CERN FAQ LHC: The Guide 4, 15-19 (2008), available at http://cdsmedia.cern.ch/img/CERN-Brochure-2008-001-Eng.pdf.

[10] CERN, All timelines | CERN timelines, http://timeline.web.cern.ch/timelines.

[11] CERN, CERN FAQ LHC The Guide at 19, http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf.

[12] Travis Lupick, B.C. Scientists Aim to Unlock Secrets of Universe, The Georgia Straight, May 15, 2008, http://www.straight.com/article-145556/end-world.

[13] See CERN, Why the LHC, http://public.web.cern.ch/public/en/LHC/WhyLHC-en.html.

[14] See Jacob Aron, Elusive Higgs wins physics Nobel, shared with Englert, New Scientist (Oct. 8, 2013), http://www.newscientist.com/article/dn24365-elusive-higgs-wins-physics-nobel-shared-with-englert.html.

[15] CERN, CERN faq LHC The Guide at 3, http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf; Particle Physics Project Prioritization Panel, US Particle Physics: Scientific Opportunities, A Strategic Plan for the Next Ten Years 2 (2008), http://www.er.doe.gov/hep/files/pdfs/P5_Report%2006022008.pdf.

[16] See Fermilab, Fermilab's Tevatron, http://www.fnal.gov/pub/science/accelerator/.

coming out of a two-year "long shutdown" involving intensive maintenance.[17] During its first run, the LHC had been limited to a collision energy of 8 TeV.[18]

Early on, the LHC suffered a couple of unanticipated mishaps that caused parts of the machine to blow up.

In 2007, during a test, a design defect caused one of the magnet units to explode. In explaining the error that led to the accident, the director of Fermilab, the American particle physics laboratory that built the magnet, said, "[W]e are dumbfounded that we missed some very simple balance of forces. Not only was it missed in the engineering design but also in the four engineering reviews carried out between 1998 and 2002 before launching the construction of the magnets".[19]

Then in 2008, shortly after the LHC's launch, a faulty electrical connection caused a mishap that damaged 53 of the LHC's magnet units.[20] Although CERN's initial reports characterized the event as a "leak", Cal Tech physicist Sean Carroll offered that "'explosion' is a more accurate description".[21] Magnets were ripped out of their floor bolts and six tons of helium spewed into the tunnel in a matter of minutes.

In 2015, CERN finished two years of refitting work to allow the LHC to operate at 13 TeV, much closer to the machine's design capacity.[22]

As an effort to explore unknown realms of physics, the excitement revolves around uncertainty. In the words of Dan Tovey, a University of Sheffield physicist on the LHC's Atlas research team:

> Individually, we all have the things that we're particularly interested in; there's a variety of new physics models that could show up. But to be honest, we can't say for certain what—if anything—will show up. And the best thing that could possibly happen is that we find something that nobody has predicted at all. Something completely new and unexpected, which would set off a fresh programme of research for years to come.[23]

---

[17] Caroline Duc, Long Shutdown 1: Exciting times ahead, CERN Updates, http://home.web.cern.ch/about/updates/2013/02/long-shutdown-1-exciting-times-ahead (Feb. 8, 2013, updated Dec. 11, 2014); Discovery News, LHC Shuts Down (Temporarily), http://news.discovery.com/space/lhc-shutdown-cern-higgs-130214.htm; Jonathon Webb, LHC Smashes Energy Record with Test Collisions, BBC News (May 21, 2015), http://www.bbc.com/news/science-environment-32809636.

[18] CERN, Press Release: LHC physics data taking gets underway at new record collision energy of 8TeV (Apr. 5, 2012), http://press.web.cern.ch/press-releases/2012/04/lhc-physics-data-taking-gets-underway-new-record-collision-energy-8tev.

[19] Pier Oddone, Directors Corner: The World Stage, Fermilab Today (Apr. 3, 2007), http://www.fnal.gov/pub/today/archive/archive_2007/today07-04-03.html.

[20] Large Hadron Collider Ready to Restart, Boston.com (Nov. 20, 2009), http://www.boston.com/bigpicture/2009/11/large_hadron_collider_ready_to.html.

[21] Sean Carroll, The Particle at the End of the Universe: How the Hunt for the Higgs Boson Leads us to the Edge of a New World location 1100 of 5498 (Dutton 2012).

[22] Webb, supra note 17 (quoting physicist Dan Tovey).

[23] Id. (quoting Tovey; internal paragraph break omitted).

6.2.2 Black Holes and the Evolving Safety Rationale

The alleged risk of the LHC is that it could spawn a planet-destroying black hole. Explaining the thinking about an artificial black-hole disaster is best done in a chronological manner. That is because the safety argument has not been static; instead, it has changed over time. As one set of safety rationales has been undermined by evolving understandings in theoretical physics, experimenters have abandoned old arguments and adopted new ones.

Questions circulating in the media about whether particle colliders might produce black holes date back at least to 1999, around the time of the start-up of the an earlier experiment—the Relativistic Heavy Ion Collider (RHIC, pronounced "Rick"), located about 100 km east of New York City [12, p. 542–553].[24]

The dust-up about the RHIC led physicists in 1999 to issue an assurance that, for the foreseeable future, no particle collider would be capable of generating the energies required to form a black hole [3]. Media interest in the question then subsided. But it turned out that the physicists' safety pronouncements in 1999 were substantially mistaken. In 2001 two separate teams of theorists demonstrated that, under certain assumptions, it actually was possible to produce black holes with a present-day accelerator—the LHC, which was then under construction [4, 8].

CERN subsequently acknowledged a need for further safety assessment work [2]. The ensuing report, issued by a CERN group in 2003, concluded that accelerator-produced black holes would pose no threat since they would rapidly evaporate via a process called "Hawking radiation" [11, p. 840 et seq.].

The next year, 2004, a very well regarded scientist called the theory of Hawking radiation into question [20]. Since the scientist was a pioneer of black-hole-evaporation theory, his opinion was potentially influential.

CERN then stopped relying on Hawking radiation as a safety rationale, and the organization commissioned a new round of theoretical work on the issue [11, p. 850]. The result was a highly complex paper, released in 2008, that rested an assurance of safety on a multi-faceted approach [7]. Authored by particle physicists Michelangelo L. Mangano and Steven B. Giddings, the paper used astrophysics analysis to conclude that, under some scenarios, synthetic black holes would be able to harmlessly coexist with Earth, since they would grow too slowly to be dangerous. Under other scenarios, the paper concluded, telescope observations of certain white dwarf stars could be counted upon to rule out the dangers on an empirical basis.

Rainer Plaga, an astrophysicist, then emerged with a paper arguing that the black hole scenario could not be ruled out [15]. In the days leading up to the anticipated start up of LHC collisions, Mangano and Giddings responded to some, but not all, of Plaga's arguments [6].

---

[24] The controversy over the RHIC principally involves not black holes, but a disaster scenario involving the creation of a "strangelet," which theoretically could physically collapse the planet into a small hyperdense ball by converting all normal matter on Earth into strange matter. I discuss the RHIC more in my prior work [11, p. 829–831] [12, p. 542–553].

The chronology of the safety debate, with its pattern of confident conclusions, new revelations, and refortification with different theoretical tacks, could be seen as evidence of results-oriented research lacking academic detachedness. Indeed, in 2010, John Ellis, a top theoretical physicist for CERN who worked on a safety report seemed to confirm this. Ellis told Physics World magazine that there had been no scientific motivation for the safety reviews, calling them a "foregone conclusion".[25]

None of this, of course, means that the LHC itself is necessarily dangerous. But it does indicate that the question of the LHC's safety is more complex than it might seem at first glance.

6.3 Conceptual and Practical Problems

The LHC/black-hole question can be used to illustrate a number of problems that may occur with traditional risk assessment in the context of leading-edge scientific experimentation. To recapitulate from the introduction, they are (1) that it is difficult to rely on experts for conducting such analysis, since those experts tend to have personal connections to the experimental work being analyzed, (2) the risk assessment may require the application of uncertain science that the experiment itself is designed to illuminate, and (3) if the probability of disaster is calculated to be very low, then that probability number is rendered virtually meaningless, since the probability of error in the derivation dwarfs the derived disaster probability.

6.3.1 The Lack of Disinterested Experts

Traditional risk analysis is "distinctly expert-centered" [22, p. 123]. But when it comes to science-experiment risk, there may be a scarcity of disinterested experts. With leading-edge science experiments, the leading experts tend to be the exact same people who are involved in the experiment—either directly or indirectly.

The LHC/black-holes question is a case in point. According to CERN, "half the world's particle physicists" come to CERN for their research.[26] The other half are an

[25] See Edwin Cartlidge, Law and the end of the world, Physics World Feb. 2010 at 12–13 (quoting Ellis: "'Every time someone comes up with a new theoretical speculation about accelerator safety, it is interesting to see why that speculation does not constitute risk, but it always comes back to the cosmic-ray argument', he says. So does that mean these safety reviews are nothing more than a curiosity? 'Correct. There is no scientific motivation for these reviews. They are a foregone
(Footnote 25 continued)
conclusion, even though the community has the right to expect CERN to demonstrate the validity of the safety arguments.'").

[26] See CERN, A Global Endeavour, http://public.web.cern.ch/public/en/About/Global-en.html; see also Elizabeth Kolbert, Crash Course, The New Yorker, May 14, 2007, at 68 ("Once the collider begins operating at full power [...] nearly half the particle physicists in the world will be involved in analyzing its four-million-megabyte-per-hour stream of data.").

extended network of friends and acquaintances. Sharon Traweek, an anthropologist who did a particle-physics ethnography, describes particle physicists as forming a restrictive, cohesive community [19]. Relationships among particle physicists are highly important, and she explains that those particle physicists who do not know each other well, want to. Thus, if one wanted to find particle physicists not part of the broader circle of friends of an allegedly dangerous experiment, doing so might prove impossible.

In the case of the LHC, however, the primary risk assessment work was not done by people with mere indirect relations. Instead, the work has been done by persons employed by or having direct ties to CERN. In 2007, CERN management set up the LHC Safety Assessment Group (LSAG),[27] and each of the five members were physicists from CERN's Theory Division [5]. One of those, Mangano, co-authored the paper that served as the foundation for LSAG's final report [7, p. 1]. The other author of that key paper, Giddings, was not employed by CERN while he was working on the paper, but he was at that time anticipating a visiting position with CERN that had been previously approved [11, p. 839].

6.3.2 The Need for Uncertain Scientific Principles Under Investigation

A second problem that may occur in trying to resolve questions of the safety of novel science experiments is that a thorough traditional risk assessment might require knowledge that the experiment itself is designed to supply.

Again, the LHC/black-holes question provides examples. The theorized phenomenon of Hawking radiation—which has been used as a safety rationale for particle collisions—has not been experimentally validated [17, p. 172]. The LHC experiment itself, however, could verify Hawking radiation.

Another unanswered question from physics, which the LHC could help answer,[28] is whether there are hidden dimensions, and if so, how many exist. The LHC may create black holes, the theory goes, only if there are one or more extra dimensions [8]. The number of extra dimensions is important to the analysis, because the issue of how fast a stable black hole could grow inside the Earth is understood to depend on the number of hidden dimensions. The Giddings and Mangano paper estimated, for instance, that if there are nine or more spacetime dimensions, it would take billions of years for any black hole to grow large enough to be threatening [7]. But if we live in a 5-D reality—that is, with just one extra dimension—then under Giddings and Mangano's analysis it might only take 300,000 years until the black hole matured and ate the Earth. Again, the ultimate level of risk of black holes rests on unknown science that the LHC may or may not shed light on.

---

[27] Video recording: John Ellis, CERN Colloquium: The LHC Is Safe (Aug. 14, 2008), available at http://cdsweb.cern.ch/record/1120625 (beginning at 6 min.).

[28] See CERN, Extra Dimensions, http://press.web.cern.ch/backgrounders/extra-dimensions.

6.3.3 *The Effect of Uncertainty in Low-Probability Assessments*

The third problem with traditional risk assessment for leading-edge science experiments is that where the probability of disaster is determined to be low, then the probability number lacks robustness, since the chance of disaster described by the probability will be much lower than the chance that the probability itself is wrong [11, p. 883 et seq].

Physicists have not published any quantification of the odds of a black-hole disaster at the LHC. In spoken remarks at Oxford's Future of Humanity Institute, however, Mangano spoke[29] of probabilities of less than 10^{-40}.

That is indeed a small number. Even when the alleged harm is the annihilation of Earth, it seems perfectly safe to ignore such an infinitesimal chance. But as particle physicist Lisa Randall said, albeit in a different context, "A prediction of low risk is meaningless if the uncertainties associated with the underlying assumptions are much greater" [17, p. 181].

Philosophers Toby Ord, Rafaela Hillerbrand, and Anders Sandberg put it this way: "When an expert provides a calculation of the probability of an outcome, they are really providing the probability of the outcome occurring, given that their argument is watertight" [14, p. 1]. Reasons the argument could be defective include mathematical miscalculations, faulty data, and uncertain assumptions.

Taking, for argument's sake, that a given probability bound is low enough that it represents insignificance, the question is left begging to what extent the estimate itself is flawed. We cannot directly measure the uncertainty of the LHC safety assessment work, as doing so would require having unavailable knowledge. We can, however, borrow error rates that have been empirically determined for other forms of scientific work. One study found that as many as one in 100 articles in the life sciences may contain errors warranting retraction [14, p. 4]. Other research found that one in 10 articles in Nature and the British Medical Journal—both elite journals—have flawed statistical results [14, p. 7]. Using these statistics as a gauge of how likely it is that the LHC risk assessment is flawed, we must say that a truer view of the maximum probability of disaster at the LHC must take into account something like a one-in-10 or one-in-100 chance that the safety assessment is in error. Of course, if the safety assessment were wrong, that would not imply that the LHC is necessarily dangerous. Nonetheless, the total ceiling on risk of operating the LHC must mostly be described by the chance that the safety assessment is flawed. And that probability, even if small, is plausibly significant.

---

[29] See Video Recording: Global Catastrophic Risks Conference, Expected and Unexpected in the Exploration of the Fundamental Laws of Nature (Michelangelo L. Mangano 2008), available at http://vimeo.com/4704040 (address at University of Oxford Future of Humanity Institute) (beginning at 47 min.) (stating that the probability that a particular white dwarf of certain characteristics would have persisted despite the laws of physics being such that they would allow dangerous black-hole formation to take place at the LHC is less than 10^{-40}, and further noting that this would be the probability for the survival of one such star, and there are multiple such stars).

6.4 Rhetorical Issues

In addition to the above described problems, there is an additional wrinkle for risk assessment in the science-experiment context. Knowledge asymmetries and opportunities for selective assertions of uncertainty can be used to reframe questions of risk. This reframing can permit experiment proponents to gain a substantial rhetorical advantage, steering public discourse in a way that is helpful for experimenters.

At the heart of this issue is something of a puzzle: Considering that traditional risk analysis involves cost-benefit scrutiny and relies on calculations and statistics [22, p. 123], one might think that this mathematically based way of evaluating risk would be embraced by scientists. Yet this is not necessarily the case. In a twist, when it came to the black-hole issue, particle physicists moved away from expressing risk in quantitative terms and instead moved to speaking of risk in qualitative terms.

6.4.1 Using Pricelessness to Avoid Quantitative Analysis of Benefits

In producing quantified probabilities of harm, traditional risk assessment thrives when it is paired with quantified benefits to use as a point of comparison. Thus, one way proponents of science experiments can avoid a negative assessment in the traditional-risk-analysis mode is to frustrate the quantification of the benefits of the experiment, thus making it incomprehensible to a cost-benefit formula.

Proponents of particle experiments will generally concede that, viewed as an investment, particle physics generates "no return".[30] At first glance, this appears to be an admission against self-interest. But, as American judge and legal scholar Richard Posner pointed out in discussing RHIC safety issues, this seeming weakness is actually a source of strength: "[I]t stumps people who people who want to argue that the costs exceed the benefits" [16, p. 148].

In truth, particle physics can yield practical benefits. The world wide web was a spinoff invention from the planning phase of the LHC. And medical proton therapy is a spinoff of particle accelerator technology.[31] But these sorts of practical dividends are rare and apparently idiosyncratic. Thus, they are not very helpful as a justification for particle-physics experiments.

So, if proponents of particle physics experiments do not attempt to quantify benefits, how do they argue in favor of their expensive experiments? They provide qualitative statements. Often these statements are emotionally charged. For instance,

[30] See, e.g., Mangano, supra note 29 (beginning at 41 min.).

[31] Jeremy N.A. Matthews, Accelerators Shrink to Meet Growing Demand for Proton Therapy, Physics Today, March 2009, at 22, available at http://ptonline.aip.org/journals/doc/PHTOAD-ft/vol_62/iss_3/22_1.shtml.

physicist Stephen Hawking characterized the LHC as "vital if the human race is not to stultify and eventually die out".[32]

A recurrent theme in the non-quantified argument for particle colliders is to make a special claim of importance for particle physics over other scientific fields. An example is what Nobel laureate Steven Weinberg wrote in his book *Dreams of a Final Theory*: "The reason we give the impression that we think that elementary particle physics is more fundamental than other branches of physics is because it is, I do not know how to defend the amounts being spent on particle physics without being frank about this" [21, p. 55].

Anthropologist Sharon Traweek puts it this way [19, p. 2–3]:

> The physicists' calling is awesome: memoirs and biographies often present this corps d'elite as unique, Promethean heroes of the search for truth. …The extraordinary scale and costliness of much physics research if anything reinforces its cultural value. The great accelerators, for example, are like medieval cathedrals: free from the constraints of cost-benefit analysis.

Thus, particle physics experimentation is a project for which no practical benefit is anticipated or sought. The upside is philosophic, and if the goal is "unmasking the cosmos,"[33] it is indubitably sublime. Once the benefit has been taken out of the realm of numbers, quantitative cost-benefit analysis is rendered moot.

6.4.2 Moving Away from the "Probability Mode"

Scientists advocating large-scale leading-edge science experiments can also repel criticism by de-quantifying the discussion of risk on the other end of the cost-benefit formula—by refusing to use numerical values to discuss the chance of disaster.

Following its experience with questions about RHIC safety, physicists learned to keep quantified probabilities out of the LHC safety debate. By doing so, physicists could avoid uncomfortable questions of what constitutes an acceptable level of risk of planetary destruction. Instead, the disaster question was cast as binary: either the experiment entails a risk, or it does not.

A window into how physicists plan the public-relations side of risk assessment comes from the video of a presentation given by CERN theorist John Ellis to his colleagues at the laboratory.[34] In the discussion following the presentation, an audience member said to Ellis, "I've noticed that, very wisely, you haven't pronounced the word 'probability'."[35]

---

[32] Jon Swaine, Stephen Hawking: Large Hadron Collider vital for humanity, The Telegraph, September 9, 2008, http://www.telegraph.co.uk/news/2710348/Stephen-Hawking-Large-Hadron-Collider-vital-for-humanity.html.

[33] See Brian Greene, The Fabric of the Cosmos 22 (Alfred Knopf 2004) ("[A]s we've continued to unmask the cosmos, we've gained the intimacy that comes only from closing in on the clarity of truth. The explorations have far to go, but to many it feels as though our species is finally reaching childhood's end.").

[34] See video recording, supra note 27.

[35] See id. (beginning at 64 min.).

"Absolutely," Ellis said.

The audience member noted that when probability comes into the debate, critics can easily make an argument that the LHC is not worthwhile. That is because, no matter how small the likelihood of destruction, since the harm is so enormous, operating the LHC may be painted as a poor choice.

The lopsided balancing of costs and benefits is not the only problem created for experiment proponents by speaking in terms of probability. There is also the problem that when scientists provide a quantitative assessment in the form of a probability bound—that is, a worst-case limit on how likely a disaster could be—this ceiling on risk may be consumed by the public as if it were an estimate of the actual probability of disaster. To take a hypothetical example, a probability bound of one in a billion does not mean that an event is likely to occur once in a billion trials. It means that the risk is not more than one in a billion—even though in reality it might be far less.

Probability bounds are not actual probabilities, and thus it is fallacious to equate the two. But it is not necessarily fallacious to take probability bounds as reasonable stand-ins for probabilities. In thinking through the ultimate question of whether an experiment should be given the green light, it may be a sensible analytical step to think in terms of the worst-case scenario—which the probability bound represents. Doing so seems particularly apropos when the endeavor under consideration has only a nebulous, philosophic benefit. Yet sensible or not, such worst-case-scenario reasoning works solely against the interests of experimenters. Thus, the disincentive to produce quantified probability bounds remains.

After the flap over the RHIC's odds, CERN steered clear of assigning probabilities for LHC/black-hole risk, instead issuing unquantified statements disclaiming all risk. The Giddings and Mangano paper provided no quantified odds; rather, it made the qualitative statement that there was "no risk of any significance whatsoever" [7]. Note that this statement allows that some possibility of danger exists, but whatever quantitative extent that risk might have is veiled behind the value judgment that the risk is insignificant.

One problem with such qualitative statements is that they are susceptible to recharacterization. For instance, the LSAG report on the black-hole issue, while relying on the Giddings and Mangano paper to provide its rationale, framed its conclusion with a rosier qualitative statement, saying that LHC-generated black holes "present no conceivable danger".[36] Upon receipt of the LSAG report, CERN's permanently constituted Scientific Policy Committee went even further, saying that the LSAG report provided "proof" that the LHC was safe.[37]

There can also be coordination of rhetoric among particle physicists in an effort to help shape public debate. As reported in a 2007 New Yorker magazine article, CERN's chief science officer Jos Engelen explained that, when it comes to LHC

[36] LSAG Version 2, at p. 1.

[37] Peter Braun-Munzinger, Matteo Cavalli-Sforza, Gerard 't Hooft, Bryan Webber, & Fabi Zwirner ("CERN Scientific Policy Committee"), CERN, SPC Report on LSAG Documents 1 (2008), http://indico.cern.ch/getFile.py/access?contribId=20\&resId=0\&materialId=0\&confId=35065.

disaster scenarios, CERN officials are instructed "not to say that the probability is very small but that the probability is *zero*".[38]

Another example along these lines comes from Ellis's account of his interactions with Cambridge University's Martin Rees—Britain's Astronomer Royal and a CERN outsider. When Rees stated that the risk of the LHC causing disaster was no more than one in 50 million,[39] Ellis reached out to him.

"I … extracted from him a statement that he'd never done an estimate, and he doesn't believe there's any risk. So he's also gone over to the not-talking-about-probability mode," Ellis said. "But I'm keeping his statement in my mail until such time as this issue raises its head".[40]

Ellis said in 2008 that, since the LSAG report came out, he had seen no discussion of risk in the probability mode.[41]

6.4.3 Constructing the Quantum Straw Man

Another way in which LHC proponents re-framed the debate to advantage was to make use of a particular kind of uncertainty—quantum uncertainty—as a way to paint the black-hole question as silly.

Nima Arkani-Hamed, a particle physicist at Princeton, proffered the argument in perhaps its most colorful and memorable form to The New York Times when he explained that there was a minuscule probability "the Large Hadron Collider might make dragons that might eat us up".[42]

Engelen offered an expanded version of the argument to the New Yorker: "In quantum mechanics, there is a probability that this pen will fall through the table", Engelen said. "All of a sudden, it will be on the floor. Because it can behave as a wave, it can go through; we call that the 'tunnel effect.' If you calculate the probability that this happens, it is not identical to zero. It is a very small probability. But it never happens. I've never seen it happen. You have never seen it happen. But to the general public you make a casual remark, 'It is not identical to zero, it is very small,' and …"[43] The reporter indicates Engelen then shrugged.

---

[38] See Elizabeth Kolbert, CRASH COURSE Can a seventeen-mile-long collider unlock the universe? The New Yorker (May 14, 2007), http://www.newyorker.com/reporting/2007/05/14/070514fa_fact_kolbert.

[39] Video recording, supra note 27 (beginning at 65 min.).

[40] Id. (beginning at 66 min.).

[41] Id. (beginning at 64 min.).

[42] See Dennis Overbye, Asking a Judge to Save the World, and Maybe a Whole Lot More, New York Times, Mar. 29, 2008, at A1, available at http://www.nytimes.com/2008/03/29/science/29collider.html.

[43] Kolbert, supra note 26.

The quantum-chance-of-anything argument was embraced by the press and the blogosphere, constituting a clear public-relations victory.[44] Yet it is fallacious. The quantum dragon is a straw man. No critic of the LHC argued that the collider should be shut down because of a generic quantum-mechanics-type chance of producing a planet-eating black hole. The argument that the LHC might destroy the planet is that, under current theory, the LHC, owing to its novel characteristics, might produce black holes, and there is no good way to rule out that such a black hole would destroy the planet.

What is especially slippery about the quantum-dragons argument is that it is clothed in the language of particle physics. By incanting "quantum mechanics", the argument seems to claim some particular relevance to particle accelerators. But that is not the case. Since there is a quantum-mechanical chance of just about anything happening, the quantum-dragons-type of straw man can be applied to any debate. For instance, a person arguing for the safety of tobacco could make the same argument in response to the allegation that cigarettes cause cancer. Such a person could correctly point out that there is, quantum mechanically speaking, a chance that cigarettes will turn into tiny sticks of dynamite and explode. The argument is just as irrelevant to a debate about tobacco as it is to the LHC.

6.5 Implications for Courts

So far, I have argued that uncertainty can make science-experiment risk inapt for traditional risk assessment and that traditional modes of thinking about risk allow adroit proponents of scientific experiments to reframe questions of risk in a way that is favorable for the experimenters. Now I will discuss what this means for courts in the context of a legal challenge to an experimental program based on the allegation that it is unduly risky.

At the outset, it should be observed that debates about the safety of science that take place in the media are likely to have an effect on litigation outcomes. Of course, courts are ideally meant to be instruments of the rule of law, and public opinion has no direct relevance to deciding issues such as whether a court should use its equitable powers to halt a scientific experiment. Yet it would be naïve to say that public opinion does not carry a great deal of weight inside a courtroom. As Chief Justice William H. Rehnquist of the United States Supreme Court observed, "Judges, so long as they are relatively normal human beings, can no more escape being influenced by public

[44] See, e.g., Sharon Weinberger, Collider May End World! (Or Spit out Man Eating Dragons), Wired.com Danger Room, April 16, 2008, http://blog.wired.com/defense/2008/04/collider-may-en.html; Dennis Overbye, Gauging a Collider's Odds of Creating a Black Hole, New York Times, April 15, 2008, http://www.nytimes.com/2008/04/15/science/15risk.html. ("Besides, the random nature of quantum physics means that there is always a minuscule, but nonzero, chance of anything occurring, including that the new collider could spit out man-eating dragons."). Bloggers and blog commenters citing this argument are too numerous to cite.

opinion in the long run than can people working at other jobs" [18]. Thus, a victory won through the media in the court of public opinion becomes a substantial hurdle to obtaining an injunction, should one be merited.

Additionally, the dynamics of the debate outside the litigation context are likely to mirror debate that would take place in court. The same ways of framing questions of risk and asserting uncertainty that are persuasive to the public are likely to be persuasive to a judge as well.

Bearing this in mind, there are ways to make questions of risk from large-scale leading-edge science experiments more meaningfully susceptible to judicial review. Here I will propose two. The first is a primary, general strategy: evaluating uncertain risks in qualitative terms. Then I discuss a secondary, adjunct strategy: testing the favorable opinions of science-experiment proponents against their opinions outside the context of the controverted experiment.

6.5.1 Evaluating Uncertain Risks in Qualitative Terms

Although courts have often looked at risk in quantitative terms,[45] in the case of science experiments where uncertainty dominates, courts should evaluate risk in a qualitative sense, conducting a kind of meta-analysis that gets above the level affected by uncertainty.[46]

This approach involves framing the uncertainty in human terms—in particular, looking to the context of scientists' assessment of their own experiment as safe. The assumption here is that experts can be influenced in their judgments on risk by their cultural orientation [13]. That is to say, experts' reasoning is vulnerable to biasing influences [13, p. 1081–1082 and 1093–1094]. As legal scholar Dan Kahan and co-authors wrote, "Like members of the general public, experts are inclined to form attitudes toward risk that best express their cultural vision" [13, p. 1094].

To form a qualitative assessment of the extent to which cultural vision could affect safety assessments about the risk of experiments, courts should look at the scientists' organizational culture, community norms, group politics, and power dynamics. These human elements may provide reasons to be confident in, or skeptical of, the scientists' judgments about the risk of their own experimental programs.

As a complement to looking at the cultural context, it will likely be of use to look at the safety argument itself. This is because certain aspects of the scientists' argument may be relevant in evaluating susceptibility to bias and cultural filtering. Simple arguments for safety will be more resistant to such biases and filters. No

[45] See, e.g., Hawai'i County Green Party, 980 F. Supp. at 1168–69 (using quantified costs and probabilities in denying an application for injunction to halt launch of the Cassini space probe carrying 32.8 kg of plutonium as a power source, which plaintiffs alleged caused a cancer risk in the event of a launch accident or navigation error).

[46] What I am suggesting here is an adaptation of an approach I suggested previously for the related problem of courts dealing with scientific questions of extraordinary complexity that are opaque to laypersons [11, p. 883–907].

amount of bias or cultural filtering would, for instance, cause someone to believe 2 + 2 = 5. But the more complex the chain of reasoning involved, the more opportunity there is for judgment calls regarding what assumptions to make and what data to reference. That, in turn, means there is more room for bias and cultural influence to determine scientists' own judgment about the riskiness of their experiments.

One can imagine readily applying this sort of analysis to the LHC/black-holes case. Some factors counseling skepticism of the risk assessment offered by CERN include the involvement of people with career stakes in the outcome of the assessment, the complexity of the safety case, and the fact that the assumptions and models used in the risk analysis require the exercise of discretionary judgment.

It may seem to be a step backward in dealing with risk to intentionally embrace qualitative reasoning in lieu of quantitative analysis. But we really have no choice. As discussed above, quantitative analysis can be rendered mostly meaningless on account of its uncertain assumptions and propensity for error. Under these circumstances, quantitative sophistication has an illusory allure, while qualitative analysis, if less prepossessing at first glance, offers more in the way of integrity.

6.5.2 Testing the Opinions of Science-Experiment Proponents by Analogy to Their Opinions Outside the Context of the Controverted Experiment

Beyond generally moving to a qualitative mode to evaluate risk, it may be probative to see to what extent scientists' assessments of the safety of their own experiments are consistent with their assessments in areas outside their own sphere of self-interest.

A simple example comes from Stephen Hawking, who opined that the LHC is "absolutely safe". Yet speaking outside of the context of LHC safety, he wrote that it is plausible that current physics theory may someday be regarded "as ridiculous as a tower of turtles" [9]. Given Hawking's understanding that today's physics theory may not endure, there may be some reason to doubt his ironclad conclusions about LHC risk.

A better example, with much richer detail, is the writing that particle physicist Lisa Randall has done about the LHC/black-hole question and about risk in the financial sector.

Randall is a proponent of the LHC, saying that the risk of the accelerator generating dangerous black holes is "essentially nonexistent" [17, p. 179]. She does admit that some uncertainties are involved when it comes to the question of accelerator-

created black holes.[47] Yet Randall's bottom-line assessment is that uncertainties in the LHC/black-hole risk question can be safely ignored [17, p. 186–187]:

> Luckily for our search for understanding, we are extremely certain that the probability of producing dangerous black holes is minuscule. We don't know the precise numerical probability for a catastrophic outcome, but we don't need to because it's so negligible. Any event that won't happen even once in the lifetime of the universe can be safely ignored.

It is interesting to compare Randall's thinking about the LHC/black-hole question to how she views risk in the financial system. Randall advocates additional regulation for the financial system. Why? Randall urges that attention should be paid even to unlikely outcomes if the harm would be very large. With insight that could easily be applied to the black-hole case, she writes [17, p. 186]:

> The financial crisis happened because of events that were outside the range of what the experts had taken into account. …Virtually no one paid attention to the 'unlikely' events that precipitated the crisis. Risks that might otherwise have been apparent therefore never came up for consideration. But even unlikely events need to be considered when they can have significant enough impact. Any risk assessment is plagued by the difficulty of evaluating the risk that the underlying assumptions are incorrect. Without such estimates, any estimate becomes subject to intrinsic prejudices. On top of the calculational problems and hidden prejudices buried in these underlying assumptions, many practical policy decisions involve unknown unknowns—factors that can't be or haven't been anticipated. Sometimes we simply can't foresee the precise unlikely event that will cause trouble. This can make any prediction attempts—that will inevitably fail to factor in these unknowns—completely moot.

Similarly, Randall explains why we cannot rely on experts in the realm of finance and economics [17, p. 195]:

> After all, 'experts' told us that derivatives were a way of minimizing risk, not creating potential crises. 'Expert' economists told us that deregulation was essential to the competitiveness of American business, not to the potential downfall of the American economy…Clearly experts can be shortsighted. And experts have conflicts of interest.

Yet Randall exempts particle physicists from such fallibility. This is despite the fact that physicists can be shortsighted. Recall that physicists in 1999 prematurely declared that present-day accelerators would not have nearly enough energy to create black holes, only to be contradicted by evolving theory a couple years later. And physicists, of course, can have conflicts of interest. Recall that CERN's safety assessment was done by CERN affiliated individuals.

Notably, Randall seems to appreciate the tension in her views. She writes, "Yet despite my confidence that it was okay to rely on experts when evaluating potential risks from the LHC, I recognize the potential limitations of this strategy and don't quite know how to address them" [17, p. 195].

---

[47] See [17] at 172 ("No one really knows how to solve systems in which both quantum mechanics and gravity play an essential role. String theory is physicists' best attempt, but we don't yet understand all its implications. This means that in principle there could be a loophole. Extremely tiny black holes, which we will understand only with a theory of quantum gravity, are unlikely to behave the same way as the big black holes we derive using classical gravity. Perhaps such very tiny black holes don't decay at the rates we expect. Even this isn't a serious loophole however.") (internal paragraph break omitted).

6.6 Conclusion

We must accept the fact that it is in our nature as humans to push the frontiers of our knowledge. And we must accept as well that we can be mistake-prone and shortsighted in doing so. If the LHC does not present a real danger to humanity, another great science experiment soon might. Acknowledging our inability to tame uncertainty is the first step in being prudent and rational with such issues. Developing intelligent means of navigating uncertainty is the next step. And it's in that vein that I have sought to make a contribution with this chapter.

While the upshot of my analysis is that the courts can, indeed, do a good job in handling uncertainty in the science-experiment-risk context, a word should be said in closing about whether courts should perform this function. Without engaging in a general defense of the rule of law, I will note that when it comes to uncertainty and risk, this book points out at least two relevant desiderata. In Chap. 7, Jordan Sand takes away from the history of Edo the lesson that giving people control over their fates provides a way to live with uncertainty. And in the book's conclusion, Corinne Bieder emphasizes the key role of trust.

The courts are well-positioned to provide both trust and a meaningful sense of control. The courts supply an avenue to trust through their role of gathering and impartially weighing evidence. And the openness of the courts to hearing complaints of affected parties can provide people everywhere with a sense of control over their own destinies. In the overall analysis, it does not appear that the existence of uncertainty militates against courts engaging in questions of uncertain risk. Instead, thoughtful reflection seems to show that judicial resolution is particularly appropriate.

By proposing tools for generalist courts to deal with risk questions such as those explored here, I do not mean to imply that engaging with such issues will be easy. To the contrary, there seems little doubt that such cases will be extremely challenging. But if we can achieve so much success in reaching new heights of knowledge about the world around us, certainly we can advance our ability to make good decisions about how to do so safely.

References

1. Baram M (2009) Biotechnological research on the most dangerous pathogens: challenges for risk governance and safety management. Saf. Sci. 47(6):890-898
2. Blaizot J-P, Iliopoulos J, Madsen J, Ross GG, Sonderegger P, Specht H-J (2003) Study of potentially dangerous events during heavy-ion collisions at the LHC: report of the LHC Safety Study Group, at cover page, http://doc.cern.ch/yellowrep/2003/2003-001/p1.pdf.
3. Busza W et al. (1999) Review of "speculative disaster" scenarios at RHIC, available at http://www.bnl.gov/rhic/docs/rhicreport.pdf.
4. Dimopoulos S, Landsberg G (2001) Black holes at the large hadron collider. Phys. Rev. Lett. 87, 161602
5. Ellis J, Giudice G, Mangano M, Tkachev I, Wiedemann U (2008) Review of the safety of LHC collisions, http://arxiv.org/vc/arxiv/papers/0806/0806.3414v1.pdf

6. Giddings SB, Mangano ML (2008) Comments on claimed risk from metastable black holes, arXiv:hep-ph/0808.4087v1, CERN-PH-TH-2008-184, http://arxiv.org/abs/0808.4087v1
7. Giddings SB, Mangano ML (2008) Astrophysical implications of hypothetical stable TeV-scale black holes. Phys. Rev. D 78, 035009
8. Giddings SB, Thomas S (2001) High energy colliders as black hole factories: the end of short distance physics. Phys. Rev. D, 65(5): 056010
9. Hawking SW, Mlodinow L (2005) A briefer history of time. Bantam Dell, New York
10. Hassanzadeh S (2012) Analysis of the causes of delay in collaborative decision-making under uncertainty in pharmaceutical R&D projects. PhD thesis, Université de Toulouse
11. Johnson EE (2009) The black hole case: the injunction against the end of the world. Tenn. Law Rev. 76:819–908
12. Johnson EE (2016) Agencies and science-experiment risk. University of Illinois Law Review. 2016:527–587
13. Kahan DM, Slovic P, Braman D, Gastil J (2006) Fear of democracy: a cultural evaluation of Sunstein on risk. Harv. Law Rev. 119:1071–1109
14. Ord T, Hillerbrand R, Sandberg A (2008) Probing the improbable: methodological challenges for risks with low probabilities and high stakes
15. Plaga R (2008) On the potential catastrophic risk from metastable quantum-black holes produced at particle colliders
16. Posner RA (2004) Catastrophe: risk and response. Oxford University Press, Oxford
17. Randall L (2011) Knocking on heaven's door: how physics and scientific thinking illuminate the universe and the modern world. Harper Collins, New York
18. Rehnquist WR (1986) Constitutional law and public opinion. Suffolk Univ. Law Rev. 20:751–769
19. Traweek S (1992) Beamtimes and lifetimes: the world of high energy physicists. Harvard University Press, Cambridge
20. Unruh WG, Schützhold R (2005) Universality of the Hawking effect. Phys. Rev. D, 71, 024028
21. Weinberg S (1992) Dreams of a final theory: the scientist's search for the ultimate laws of nature. Pantheon Books, New York
22. Wilson MJW (2011) Cultural understandings of risk and the tyranny of the experts. Or. Law. Rev. 90(1):113–190

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 7
What Can Japan's Early Modern Capital of Edo Teach Us About Risk Management?

Jordan Sand

Abstract The city of Edo, early modern capital of Japan, was built of wood and burned with extraordinary frequency. This essay considers the logic of fire prevention and response in Edo in contrast to disaster management in the present day, with particular attention to the Fukushima nuclear disaster. The comparison reveals that the emphasis in Edo on strength and continuity of the social order rather than preservation of material property produced a different view of risk and uncertainty.

Keywords Urban fire · Japan · Fukushima nuclear disaster · Edo-Tokyo · Psychology of risk

7.1 Introduction

The city of Edo, precursor of present-day Tokyo, was the capital of the Tokugawa shogunate, a dynastic military regime, from 1600 to 1868. Under the quasi-feudal system set up by the Tokugawa family, the 250-odd provincial lords known as daimyo, who ruled independent domains throughout the country, were required to live half their time in the capital and maintain large permanent estates with bands of retainers there. These military men, called *bushi* or *samurai*, made up roughly half the city's population. Servicing this large ruling elite was the foundation of the city's prosperity.[1] Since the Tokugawa shogunate limited foreign trade to the far western port of Nagasaki and allowed no foreign residence elsewhere in Japan, the growth of international trade and Western imperialism had little direct impact on the city until the arrival of U.S. gunboats in 1853. The country also remained largely free of internal warfare during these two centuries. Despite the bureaucratization of rule over the course of this long peace, the Tokugawa rulers remained military men, whose

[1] *Bushi* is the usual term in Japanese for all members of the military caste, who made up roughly six percent of the national population.

J. Sand (✉)
Georgetown University, Washington, D.C, USA
e-mail: sandj@georgetown.edu

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_7

Fig. 7.1 Map of fires in Edo-Tokyo

capital was in theory, if not in practice, a fortress and encampment. The merchant bourgeoisie that came to flourish in Edo was given considerable autonomy but was never involved in the governance of the city as a whole.

Since almost all construction in this densely populated metropolis was of wood and the technology of fire extinguishing was primitive, large fires were routine.

This map (Fig. 7.1) by Yamakawa Kenjirō, Tokyo University's first native professor of physics, shows 91 fires that burned 15 blocks or more over a period of 234 years. The map was published in 1888, the year after the Meiji government completed the first comprehensive fireproofing project in central Tokyo. Yamakawa shows the point of each fire's origin and its point of extinction, in order to illustrate the influence of the prevailing winds each month of the year.[2]

Once or twice each generation, a fire would cause high casualties. Historian Kuroki Takashi records 16 fires in the period from 1601 to 1859 that caused over 100 deaths, of which eight caused over 1000 deaths. Such high casualty figures are shocking by contemporary standards for fires. But how did life in Edo compare in actuarial terms to life in other cities? Based on the numbers of dead and missing reported by Kuroki for the period between 1755 and 1855 (a period that includes the Ansei earthquake of 1855, but excludes the anomalous and statistically dubious Meireki fire of 1657), and assuming a constant urban population of one million, the average

[2] Yamakawa Kenjirō, "Tōkyō fuka kasairoku," *Rika kaisui 3:2, Tōkyō kishōhen*, 1881.

annual rate of deaths due to conflagration comes to roughly $\frac{32}{100000}$.[3] This would place the risk of death from fire in Edo in the same range as the contemporary risk of being murdered in Oakland, California, which ranked number 7 in the list of the top ten US cities for homicide rates in 2012. Since the data for Edo is fragmentary, however, the estimate based on Kuroki's figures may be low. Doubling it would place Edo's fire risk alongside the homicide risk in Flint, Michigan, the "deadliest city" in the United States in 2012.[4] It would appear, therefore, that Edo was a risky place, but people live in riskier ones, possibly even in major American cities today. On the other hand, one was considerably more likely to die from epidemic disease in Edo than to die in a fire, and in this respect Edo was a safer place than most cities in Europe at the time. Before cholera reached Edo in 1858, deaths due to epidemic disease were higher as a percentage of population in Europe than they were in Japan [6, pp. 188–208]. Overall, life expectancy in 18th-century Edo probably compared favorably with London or Paris.

The real question to consider in thinking about how fire risk was managed, however, is how people perceived the risk. In this regard it is important that only a small percentage of the fires that turned into conflagrations destroying multiple houses in Edo also resulted in deaths. Disaster management scientist Nishida Yukio calculates that roughly 5 % of recorded fires in the second half of the Edo period resulted in deaths [18]. This meant that Edo residents frequently witnessed large fires, and were even forced to flee from them, without experiencing a direct threat to their own lives.

The issues raised by Edo's fire problem thus go beyond statistics. This study will draw comparisons to present-day disasters in order to consider what fire management in Edo reveals about the social mechanisms and social psychology of disaster preparation and response. I will refer particularly to the case of the Fukushima nuclear accident, which presents a striking contrast and suggests the extent to which Japanese society was transformed in the twentieth century.

My claim that the people of Edo may have something to teach us is premised on a few anti-orientalist assumptions about early modern Japan. The first is that Edoites were not fundamentally different from ourselves. They valued their lives and took precautions to protect them. They were not fatalists who accepted all misfortunes as inevitable. Nor did they worship their leaders. Edoites belonged to a largely literate urban society. Although the authorities censored print and other media heavily, criticism of government failures can be read between the lines of a vibrant tradition of social satire. Second, Edo was not a static society. Although institutional and cultural change came slowly under Tokugawa feudal rule, Edo saw significant economic and technological development, making it comparable to early modern Europe in many respects. Techniques were developed to battle fire and authorities were aware of other systems of fire hazard reduction. Eventually, following the Meiji Restoration

[3] Kuroki Takashi, *Edo no kaji* (Tokyo: Dōseisha, 1999), 229–235. The total of Kuroki's figures for the period is 31, 860. This omits several large fires about which the historical record indicates only "numerous deaths".

[4] Lauren Galik, "America's Ten Deadliest Cities 2012", *policymic* January 11, 2013. policymic.com/articles/22686/america-s-10-deadliest-cities-2012.

of 1868, the government embarked on the road toward the chimerical goal of a zero-hazard environment. For two centuries prior to this, no such goal was pursued. Yet as the city's foremost environmental hazard, fire prompted a range of responses and innovations.

The first and most obvious contrast with the present day is that in addition to being feared and lamented, fires in Edo were also celebrated. A popular saying went that "fires and fights are the flowers of Edo", suggesting that both were enjoyed as spectacles.[5] Fires followed certain rhythms, coming most often in the dry winter months, becoming conflagrations when the wind was high (usually out of the northwest), and spreading at a rate that permitted warnings to be relayed ahead of their arrival. This familiarity reduced the trauma of the experience for survivors. We see similar examples in other situations. The Bangkok floods of 2011, for example, were reported in the world press as catastrophic in proportion and causing property damage in the millions, yet there were also reports of Thai residents going about their business apparently cheerfully in streets inundated with over a meter of water. Flooding is a familiar part of life for residents of much of the region around Bangkok. They are more than simply resilient—they have accommodated their lives in practical terms to the hazard. I should emphasize that this does not mean that we should consider the residents of these disaster-prone places as content and therefore needing no remedy. It means rather that familiarity can make people capable of extraordinary adaptations and that the language of trauma may not always be appropriate for interpreting the social effects of a fire or flood.

Every type of disaster traces a distinctive temporal arc. The entire arc of a large conflagration in Edo, from first warning until first steps of recovery, is presented in excellent detail in the narrative scroll painting titled "Views of an Accidental Fire and its Extinguishing in Edo", painted by artist Umesawa Seiga in 1829 and presumably representing the colossal fire of that year, which damaged or destroyed tens of thousands of houses.[6] The scroll (shown in Figs. 7.2, 7.3 and 7.4[7]) begins with a fire watchtower at a daimyo estate. Further along, commoner townsmen appear, some in a fire lookout on a rooftop, while across the street others pack up belongings, suspending chests of drawers from poles. This portable furniture represents one of the key adaptations Edoites had made to fire. They lived constantly at the ready to flee with their belongings. Nearby are several groups of people who have fled and are taking refuge in a wide area, probably a firebreak. They have formed private enclosures out of furniture, doors and tatami floor mats from their houses. Edoites were accustomed to arranging temporary shelter for themselves in this manner. Men on a rooftop nearby hold enormous hand fans to fan away flying embers. Then comes the fire itself, and firemen jumping the flames on a rooftop. Rather than extinguishing

[5] For discussion of the culture of Edo fire brigades, see [8].

[6] *Edo shikka shōbō no kei*; Tachibana Museum, Yanagawa. Reproduced in *Ō-Edo happyakuya chō* (Edo-Tōkyō hakubutsukan, 2003), 80-83.

[7] This figure and the two following figures are used with the permission of the Tachibana Museum, Japan, where the originals of the scrolls are exposed. These three figures are not covered by our Creative Commons Attribution licence.

Fig. 7.2 Residents gather their belongings while firemen run by

Fig. 7.3 Refugees set up temporary households with their doors and belongings in a clearing

the flames, the firemen strip off the roof to allow the energy of the fire to exhaust itself and to limit its lateral spread (many firemen were in fact also roofers). Then we see the conflagration itself, followed by a burned-out neighborhood where the fire has already passed. The only buildings still standing are clay-walled storehouses, sealed and labeled. Some are being doused. Firemen clean up, while gleaners can be seen sifting through the ashes searching for nails and other pieces of metal to be recycled. Already a fence is being erected around one site, revealing how rapidly shopowners began recovery so as to lose as little custom as possible. The point at which the fire stopped can also be seen: a partially burned house festooned with the placards of the fire companies. The companies would be compensated based on this evidence of their participation in firefighting. Elsewhere, shogunal officials inspect the damage. Householders whose houses survived reinstall tatami floor mats. People go about greeting one another, bowing, conveying thanks and condolences. Finally, the painter celebrates the firemen's bravery by showing them performing acrobatics on a ladder—a performance still seen in Tokyo today at New Year. The acrobatic display reminds us that in Edo firefighting, technique and daring took precedence over mechanical technology. Finally, two samurai are seen greeting each other in the entrance of an unharmed estate, suggesting a return to peace with solemnly performed acts of social exchange. Their grooms wait outside, regaling each other with fire stories.

Sadly, a sequence of vignettes like this, tracing a coherent trajectory from disaster occurrence to resolution and recovery, is impossible to imagine for the Fukushima nuclear disaster. Not only is the return to normal conditions in Fukushima projected for a distant future, the response of authorities has been dogged by doubts and scandal and the affected population lives in a state of profound uncertainty. It is ironic, in a sense, that strictly in terms of loss of life, the 1829 fire caused greater casualties

Fig. 7.4 After the fire has passed, people reinstall floor mats and doors, and pay condolence visits to neighbors. In the foreground, the charred roof beams of a house that was stripped by firefighters

than Fukushima is expected to do by present estimates. But even this is uncertain: we won't know for years how many lives will be shortened by the radiation. Although present projections of casualties are low, the question will probably always remain ambiguous. What is certain is that the Fukushima accident, unlike a fire in Edo, is open-ended. It has precipitated wide-ranging confusion and a state of uncertainty that may last for generations.

7.2 Six Principles Drawn from Edo's Fire Management

The remainder of this chapter will propose six principles derived from the study of fire and its management in Edo for considering risk and disaster in the modern world. We should not look at history with rose-colored glasses. The Tokugawa shogunate was an autocratic regime that imposed a hereditary status hierarchy to preserve the privilege of a few against the many. The regime's treatment of criminals and suspected criminals was draconian. Most Edoites lived in poverty with little chance of improving their lot. Tokugawa society was not one in which many of us today would wish to live. My intent here, therefore, is not to propose a return to Tokugawa-period practices. Nor do I intend to champion Japanese culture's supposedly inherent virtues in times of adversity, as was common in the mass media after March 11, 2011. The aim is rather to use the case of Edo's fires as a means to reexamine values and assumptions in today's disaster and risk management discourse.

7.2.1 Build Light, Travel Light

At the most basic level, this principle is an obvious one. The more we invest in technology, buildings and material goods, the more we have to lose. Lighter structures provide less fuel for a fire, and less rubble to kill, injure or trap occupants in a collapse. Lightweight, portable belongings can be carried to safety, easing the return to normal life after a disaster. Terrible dangers accompany this portability too, however, since refugees encumbered by baggage may become a hazard to themselves and others. In Edo's first and worst conflagration, the Meireki Fire of 1657, thousand of lives were lost because residents attempted to rescue their belongings in large four-wheeled

carts, which caused traffic jams and provided further fuel for the flames. The shogunate subsequently forbade the use of four-wheeled carts [11, p. 115–116]. As we see in Umesawa Seiga's painting, household belongings thereafter were carried on the back, suspended from poles or, occasionally, placed on maneuverable two-wheeled carts.

Edo's buildings too were designed to be dismantled and partly portable. There were three classifications of building types in Edo: clay-walled storehouses, plastered wood structures with tiled roofs, and tenements built entirely of wood, with wood shingled roofs. In a large fire, the storehouses were sealed, plastered wood structures were stripped to their structural skeletons, and wooden tenements were treated as expendable and torn down in the path of the fire to diminish the available flammable material and the risk of flying embers. Residents carted off doors and tatami floor mats. Since these pieces were modular, they could be immediately installed when houses were reconstructed. The same modularity made it possible to dismantle, transport and reassemble the buildings themselves in non-emergency conditions.

The ultimate implications of the "build light, travel light" ethos are ecological. A city is the most massive object constructed by human effort. Hypothetically, it is possible to imagine the total weight of all the steel, stone, brick and mortar, infrastructure and material goods accumulated in a city. Learning from the ecologist's term "biomass", we could call this totality "technomass". What would a comparison of the technomass of different cities, or of cities in different eras reveal? Or of the biomass/technomass ratio in different cities? This kind of hypothetical calculation also has bearing on issues of environmental sustainability. A calculation of technomass would tell us something about how much energy input was required to erect the city, or to reconstruct it after a disaster. For its population size, the city of Edo was at the light end of the technomass spectrum.

A nuclear power plant would stand at the heaviest end of the spectrum. Although the hazards of a nuclear plant lie partly in the intrinsic character of nuclear fuel and partly in the rigidity of the system for controlling its fission, which allows small margins for error, the Fukushima disaster demonstrated that the problems of nuclear power are compounded by the sheer mass of the technology, which must be heavy to be preserved in place for an indefinitely long time, protected from seismic and other environmental disruptions. Seismicity poses a special problem for Japan, one of the most seismically active inhabited areas on the planet. The trade-off between fire resistance and earthquake resistance was first seriously confronted in Japan following the Nobi Earthquake of 1891, in which brick structures that had been built as demonstrations of the modern state's technological progress collapsed, causing heightened casualties.[8] The government had initiated a heavy-technomass approach to urbanism two decades earlier with reconstruction in brick of buildings on the main street of Tokyo's Ginza commercial district following destruction of the area by a fire. The response after 1891 was not to return to lighter construction, however. Permanent structures were viewed as essential to capitalist modernity. Eventually, reinforced concrete would emerge as the solution, together with heavier and heavier

[8]On the engineering debates sparked by the Nobi Earthquake, see [2].

infrastructure. In broad terms, the Fukushima nuclear plant, constructed in 1970, can be seen along a single trajectory from the brick-built, gas-lit Ginza of 1872.

7.2.2 Strong Social Infrastructure Counterbalances Vulnerable Physical Infrastructure

Here too, the basic principle is an obvious one: mutual aid helps victims overcome the effects of a disaster. By "social infrastructure" I mean the network of relationships and social institutions that support urban life and provide assistance in times of need. In a city like Edo, where the physical infrastructure of the city was fragile and easily destroyed, residents were likely to need the assistance of neighbors at some point in their lives. The fire hazard encouraged mutual aid and mutual aid made the fire hazard more manageable.

There is a tendency to idealize strong social infrastructure today because it is in short supply in first-world settings. In Japan, communalism was often coerced. The people of Edo were required to police themselves and field their own fire brigades. The wealthy property owners who had the most to lose from lawlessness and property damage paid directly for these services, which were provided by their tenants and employees, and established town residents provided alms to the poor until houses were rebuilt. Although this charity was in theory voluntary, the need to maintain neighbors' respect (and keep their custom) made it effectively obligatory [9, p. 230]. The modern government exploited these traditions, particularly in the era of fascism and total war. While American bombers showered incendiaries on downtown Tokyo, citizens were urged to stay and run bucket relays, which were almost completely ineffective. They participated in this futile effort as members of neighborhood associations, whose strength was a legacy of local organizations dating back to the Edo period [5, pp. 155–160].

Strong social infrastructure in Edo was thus the product of both environmental pressures and political pressure from above. The Tokugawa shogunate treated fires in the commoners' districts predominantly as a problem of social control, and addressed them through moral exhortation and threat of punishment. Each block (*chō*) was required to have a guardhouse with a watchtower, and guards, accompanied by a local concierge, were expected to make nightly rounds. Commoners were subject to punishment for accidentally starting a fire that spread to adjacent buildings. To encourage mutual surveillance, punishment was also meted out to the landlord, concierge, a neighborhood official responsible for the fire watch, and members of the five-household group that made up the smallest unit of urban administration [11, pp. 130–131]. The communal responsibility for fire control enforced the maintenance of a strong social infrastructure.

The kind of mutual dependence among neighbors that characterized urban Japanese life before the modern era is probably impossible to recover in first-world cities of the present day, even if we were to desire it. Signs of mutual aid and sto-

icism in Japan after the March 11th, 2011 disaster were widely reported and praised. The absence of looting was taken as evidence of social cohesion. Certainly some of the hallmarks of traditional social infrastructure and resilience are intact in Japan. Yet in modern Japan as in other advanced nations, citizenship has been channeled through state-dominated institutions, making the majority of citizens passive—if largely law-abiding—participants in society. In this context, it is difficult to imagine the type of disaster response seen in Edo occurring today. Events in the affected areas in northeast Japan since the 2011 disaster are instructive in this connection. A few tightly knit villages outside the radiation zone revealed the strength of traditional community by rebuilding together, but more commonly, large construction interests tied to the ruling Liberal Democratic Party have taken over the process, often without regard to the future needs of the region. Meanwhile, thousands of Fukushima refugees remain stranded in a bureaucratic limbo, living in government-built temporary housing, waiting for compensation, unable to rebuild their lives.

7.2.3 Disaster Clarifies Society's Hierarchy of Values

Edo's frequent conflagrations appear from a modern perspective not only a risk to life and limb, but a massive and wasteful destruction of property. One might imagine the economic impact to have been devastating. Yet Edo flourished as a commercial city, and following the Meiji Restoration, transformed with extraordinary rapidity into the capital of an industrial country. The Tokugawa bourgeoisie had managed to build an advanced commercial society without depending on accumulation of fixed capital in the city. The shogun's government generally tolerated but did nothing to encourage this commercial development or to protect private investment. The hierarchy of things protected or saved in the case of fire gives testimony to these values. The shogunate placed priority on protection of the shogun's palace, expecting daimyo and townsmen to fend for themselves. As for the commoners, clearly the priority was placed on life and limb rather than the protection of property, particularly of buildings. This might seem obvious and universal, yet the logic of private property in many societies today has distorted this value, producing disaster responses that would have puzzled an Edoite. In Malibu, California, for example, firefighters in municipal and state brigades are dispatched at great public expense and individual risk to quench naturally occurring brush fires that threaten the mansions of wealthy citizens who have chosen to build in a hazardous environment. It would appear that the state of California values private wealth—or the unconditional promise of aid to all private property holders, no matter how wealthy—to such an extent that it will place public servants' lives in danger to save rich people's houses.[9]

The Vehicle of Fire Precautions, a Japanese advice manual published in 1766, instructed readers facing an approaching fire to attend to their affairs in the following order: household members, contents of the family shrine or Buddhist altar, cash and

[9]On the politics of firefighting and property in Southern California, see [3].

goods, then storehouses and cellars. The text advised first that the family should put on fine clothes before fleeing, since it was advantageous not to appear destitute when calling on strangers for aid. Planning for the period of uncertainty during and following the disaster itself thus took precedence over preservation of any material property. The next priority was placed on saving religious objects associated with ancestors, which embodied the family's spiritual continuity even if all other property were lost. Following portable wealth, buildings came last, in the form of storehouses and cellars for preserving portable goods and cash. No mention was made of attempting to preserve the house itself [12].

It is not impossible to imagine a readjustment of the contemporary hierarchy of disaster response priorities in the direction of the Edo approach, in which the government offers no guarantee for the security of private property. Faced with the possible loss of everything, people today might also make similar choices to those made by Edoites. The victims of the 2011 tsunami in northeast Japan had undoubtedly invested more in their homes than the average Edoite. Solid and secure new housing thus plays an important role in reconstruction planning. Yet at a psychological level, one of the most striking things to emerge in the massive recovery effort has been the importance of family photographs. Tens of thousands of photographs were collected from the debris left by the tsunami, and both private organizations and government bodies contributed to cleaning them, exhibiting them, and seeking to locate their owners.[10] A personal photograph or family album generally belongs outside the category of insured property, indicating that its meaning as property has not been translated into monetary terms (in this sense, the insurance industry, which depends on a commodification of things that are often uncommodified, also distorts value). Yet like the religious articles of family worship for the Edo resident, the family photograph may have a symbolic value that exceeds the price put on other material things.

7.2.4 Give People Some Control Over Their Fates and They Will Tolerate Risk

As suggested above, the contrast between Edo fire victims and victims of the Fukushima nuclear disaster appears stark. Known casualties for Edo's largest fires reached into the thousands. At present, no one is known to have died due to radiation from the Fukushima nuclear disaster. Despite this disparity, the Fukushima disaster's impact on residents of the region and beyond appears more traumatic than the impact of most fires on Edoites. Caution is due in the comparison of emotional effects, since the information about Fukushima victims that is provided to us by contemporary media is quantitatively and qualitatively different from what comes down to us in the historical record of Edo's disasters. Certainly Edoites grieved, and descriptions

[10]David Slater, "Incurring Debt; Picturing Death: Japanese Family Albums Washed Away in Tsunami Waters," Forthcoming.

survive that witness with horror the deaths of helpless citizens. Yet the social psychology of the overall disaster response in the two cases contrasts sharply, revealing the very different nature of the two types of disaster. I believe the essence of the psychological contrast lies in knowability and the sense of control.

Despite the high casualties that occasionally attended a conflagration in Edo, a fire was a knowable disaster. It was spectacularly visible. Experience and close attention to wind direction permitted one to judge a course of action (and to dress one's children in fine clothes in case flight should prove necessary). Edoites in a fire were like highway drivers today, who know the danger of driving but seldom fear for their lives because control of their own mobility affords them the illusion that they are safe. Since the fire brigades worked primarily by demolishing or dismantling houses rather than extinguishing the fire with water, and since the majority of houses were rented, most citizens packed up what could be carried and left rather than waiting for rescue. Once they had found themselves a place they believed safe, they could join the throngs watching the fire, which was a popular sport. This pastime was of course nothing peculiar to Edo: the sublime power of nature manifested in a fire seems to have an innate attraction. Repeated edicts from the city magistrate's office reveal the frustration of the authorities with gaping crowds who got in the way of firefighters' work [11]. American Edward Morse, who lived in Tokyo in the 1870s, reported finding on the way home from an evening's fire-viewing at one o'clock in the morning that a number of shops had opened their doors to pick up spectator traffic [17].

On a day-to-day basis, Edoites were aware of a list of precautions to take against the fire hazard, including avoiding the use of candles on the second floor of houses (where they were more likely to be left unattended and forgotten), maintaining water in cisterns and buckets at all times, dispatching men on regular fire watches during winter nights, and sealing up shutters on windy days. When a fire was burning elsewhere in town, residents were enjoined to climb on the roofs of their houses to extinguish or fan away flying embers. These injunctions were enforced by the shogunate, transmitted down to authorities at the district and block level, and repeated in manuals like the *Vehicle of Fire Prevention*. A fire usually began as a private accident, in someone's home, making it a personal responsibility. The great majority were extinguished on the spot. Even when it became a conflagration, the origin could be known and was reported. Writing in his diary on the 15th of the first month of 1846 about a fire that had begun two hours away from where he lived, Kanda district headman Saitō Gesshin recorded the address and the name of the head of household at the house where the fire had started.[11] Thus, despite the reality that uncontrolled fire overwhelmed human capacities and swept through large swathes of the city periodically, Edoites were able to understand a fire as a preventable calamity and one over which the individual might exercise some control. Even after it had engulfed blocks, with just a little luck one could be sure at least of one's own bodily safety.

The Fukushima disaster presents the opposite pole in terms of knowability and control. Its effects are invisible, its nature and scope incomprehensible, and its victims

[11] Quoted in [13].

completely powerless against it. This is a dilemma of contemporary society as a whole. As Ulrich Beck writes, "non-knowledge rules in the world risk society" [1]. In place of an immediate, visceral knowability, we have unknowability produced by science itself and batteries of scientists speculating on what cannot be known. Although some specialists in Japan have argued that the risk posed to individuals by radiation from Fukushima was in fact minute, the incident has bred an extreme sense of vulnerability, together with distrust and resentment. The seemingly simple fact of what can be apprehended visually expresses the disparity in social psychology. In the first days after the tsunami, the Fukushima nuclear accident unfolded on television as if in slow motion. Cameras stayed fixed on the plant although most of the time nothing appeared to be happening. Hydrogen explosions occurred on the second, fourth, and fifth days. These provided the only visual evidence that a disaster was occurring. As a result, they were televised repeatedly. As time passed and reports of the radiation leaks became more serious, no new imagery evoked the disaster in spectacular form. Instead, intimate images underscored individual helplessness in the face of an invisible threat: people carrying geiger counters (even school children—although there was little they could do if the radiation levels rose), medics in protective suits conducting radiation tests on residents who were wearing no special gear, closed roads to settlements that looked untouched but to which residents were forbidden to return.

The difference between early modern Edo and present-day Fukushima with regard to knowability and the sense of personal control in the face of disaster correlates with divergent understandings of how much risk is tolerable. Edoites lived with a threat to their lives that was constantly present, but it was familiar, visible, and knowable. They developed a range of individual and collective measures for prevention and response. These habituated practices helped give the threat a finite form and shape. A nuclear accident offers little of this. It is not even clear where the limits of the affected region and population lie. The desire to give the threat a containable form is not diminished by these frustrating conditions, however. Some people as far away as the Pacific coast of the United States took iodine pills in the hope of protecting themselves from the risk of thyroid cancer resulting from Fukushima radiation. From a scientific perspective, the probability that this was necessary or useful must have been infinitesimally small, yet those who did it could at least feel that they were responding actively. Like carrying a geiger counter, it represented taking some control of the condition of risk.

One common point between Edo fires and the Fukushima nuclear accident is the affixing of blame. In both cases, people sought to identify a person or group of people to be held responsible. This is evident in the shogunate's threats of punishment against both individuals and the neighborhood leaders and watchmen who allowed an accidental fire to start. A more extreme manifestation appears in the hunt for arsonists. During a violent crackdown on arson in 1722 and 1723, Shogun Yoshimune had 101 accused arsonists burned at the stake. Nearly half of this number were outcastes or drifters. Another 226 outcastes were banished to remote islands for the crime of arson [14]. This form of arbitrary justice kept the status hierarchy in place while at the same time making fire a containable event to be dealt with by punishment, rather

than a chronic dilemma of urban management. Among the common people, the head of a household where an accidental fire had occurred accepted responsibility for the damage or threat by going around to all neighbors with gifts and apologies. In fact, a fire produced extended rounds of gift-giving, as relationships of trust were shored up by acknowledgments of local responsibility and condolences to those affected. All of these practices sought to reestablish the social infrastructure on a secure footing. In attenuated form, they survive in Japanese communities today.

The Japanese public gave the leadership of the Tokyo Electric Company (Tepco) and the government of Prime Minister Kan Naoto a large portion of the blame for the Fukushima accident and for its mishandling. I will not here attempt to assess whether this was justified or unfair scapegoating. However, I think we can see it in general terms as part of the understanding of a disaster—any disaster—in a modern democratic society. It may be juxtaposed with Yoshimune's execution of accused arsonists, which stands as its autocratic counterpart. Tokugawa autocracy revealed its effectuality through arbitrary displays of the shogun's control over life and death, in this instance, by using execution by fire to punish accused arsonists, with a symmetry that matched the punishment to the alleged crime. In modern democratic society, we expect leadership to be manifested in demonstration of transparency in institutional decisions and personal acceptance of the burden for institutional failures. Both responses create legibility out of confusion by isolating individuals and emphasizing their malfeasance above other factors.

All acts of fixing individual or group responsibility can thus be seen as products of the impulse to make the disaster knowable and to regain control. Generally speaking, they have a conservative political effect, righting the ship of state and returning society to business as usual. By the same token, they hinder the ability to view the causes of the disaster systemically and to recognize the spectrum of social contradictions in which a broader perspective would compel us to situate it. The Tokugawa shogunate never took it upon itself to prevent conflagration and protect the welfare of all citizens, including the outcastes and others who suffered arbitrary justice, since such ideals conflicted with its model of a hierarchical society. Only a minority of those criticizing the Kan government and Tepco after 2011 asked whether part of the cause for the disaster might lie in the lifestyle they themselves had come to take for granted, which depended on power delivered from Fukushima.

7.2.5 Personal Risk Differs from Systemic Uncertainty

Whether or not the fire hazard in Edo presented an acceptable risk is a question that submits theoretically to ordinary judgment, since the risk can be calculated numerically, at least in approximate terms. Edo citizens would not have known these statistics themselves, but both folk transmission and published books would have taught them the litany of past disasters, and many would have seen casualties in fires during their own lifetime. Awareness of this kind of personal risk does not necessarily determine one's outlook on life as a whole, however. I suspect that more residents

of an economically depressed American city with a high homicide rate today have a bleak view of the future than did residents of Edo, at least in times of relative prosperity, as the 1810–1820s were reported to be, for example. Edo enjoyed social and political stability from the middle of the seventeenth century until the middle of the nineteenth century. Two hundred years of comparative security, when combined with the experience of frequent fires and subsequent reconstruction, allowed Edoites to hold the expectation after each fire that life would soon return to normal. For men in the building trades, a large fire signaled boom times, since there would be lots of work in the reconstruction and wages would rise. Since frequent demand made these men a large part of the work force, the increase in their wages had a spillover effect on other businesses too.

All of this sustained a view of periodic destruction and reconstruction as a cycle with positive effects as well as negative ones. The general understanding was that eruptions of natural violence released pent-up force and kept both nature and human society healthy by renewing them. Confucian philosophers as well as ordinary people believed that the economy followed the same principle. Just as the vital force known as *ki* (Chinese *qi*) flowed continuously in nature, money should be kept moving in the economy too, not allowed to stagnate and foster greed [15]. For this reason, many people viewed capital accumulation distrustfully. Nature, they believed, censured it.

One manifestation of this thinking can be seen in the outpouring of satirical woodblock prints during the two months after the massive Ansei earthquake, which destroyed much of the city and cost between four thousand and seven thousand lives in 1855. Based on a folk belief that earthquakes were caused by a giant catfish, the prints depicted catfish in a variety of guises: wreaking havoc, being subdued by gods or common people, even dressed as firemen saving people from the wreckage. Many of the prints lampooned the wealthy, whose shops and storehouses had collapsed, and depicted the celebratory mood among the carpenters and roofers who stood to profit. A few made oblique reference to the troubles of the Tokugawa regime, which was faced with a crisis precipitated by the recent arrival of U.S. gunboats demanding an opening of Japan's ports. The prints were soon banned [16, 20].

It appears improbable enough that a disaster of almost unprecedented scale could inspire comedy. No less surprising is the fact that many of the prints referred to the present with a phrase that meant "this blessed age," or "this blessed reign" (*arigataki goyo*). The words probably mixed equal parts irony and genuine hope. Although Edo commoners were kept out of national politics, they knew that the political elite were in an unprecedented position of danger because of the American incursion. The earthquake disaster threatened to destabilize conditions further. Yet disasters had visited the city repeatedly in the past, and each time, the same cycle of destruction and renewal had been repeated. As in the past, the shogunate had town leaders draw up records of alms contributions, obligating rich and poor alike to give toward the welfare of all in rough proportion to their wealth. Rebuilding presaged one or two years of high wages. Formulaic reference to a "blessed reign" may also have reflected the hope that political stability would return along with prosperity [10, p. 235].

Modern nation-states claim to assure the safety and welfare of all citizens (which is not to say that any state truly guarantees the safety of citizens or distributes risk

equitably). Yet when a disaster—whether it be war, natural catastrophe, or industrial accident—throws into question the long-term viability of the system itself, revealing that the life we presently enjoy is probably unsustainable, we confront a kind of uncertainty that differs fundamentally from that engendered by a familiar hazard. This is true even when calculations can be made that indicate individual endangerment to be low in the short term. Thus awareness of uncertainty may have no correlation with quantifiable individual levels of risk.

7.2.6 Safety Decisions Are Political Decisions

This principle is discussed at length by Mary Douglas and Aaron Wildavsky in *Risk and Culture*. Stated as I have done here, it is something of a truism. Different people will be affected by a disaster differently, and society's resources to cope with hazards are finite, meaning political choices must inevitably be made. "Is it fair?" may be as important a question as "Is it safe?" But as Douglas and Wildavsky emphasize, we need to recognize not only that choices have to be made, but that particular political models underlie our objectives and priorities in coping with safety issues. The case of Edo becomes heuristically useful in this context because its founding political assumptions contrasted so sharply with the dominant assumptions today [4].

Although the particulars of economic and social policy fluctuated over the long period of Tokugawa rule and some members of the merchant bourgeoisie became rich far beyond their station, the shogunate remained consistent in maintaining the formal structure of status hierarchy. Land held by samurai, commoners, and religious institutions was segregated. The samurai estates that occupied seventy percent of the area of the city were considered to be granted temporarily by the regime, and not to be traded. In commoner districts, hereditary town elders bore responsibility for all in their districts, maintaining curfews and acting as local proxies for town magistrates. Landlords and their concierges had authority over their tenants approaching that of a parent under modern law. As mentioned earlier, collective punishment was occasionally used to remind all parties that responsibility accompanied authority.

This hierarchical ideal, rather than the goal of either universal welfare or prosperity, ran through the government handling of fire prevention and response. The historical exception proves the rule. In the 1730s, the unusually activist shogun Yoshimune (the same Yoshimune who had burned arsonists at the stake) sought to make the capital city more resistant to fire. He offered cash loans to samurai and relief from tax and corvée to commoners in order to tile their roofs and plaster their walls. Leaders of commoner districts petitioned the shogunate, complaining of prohibitive cost. Some textual evidence suggests that substantial fireproofing was achieved, but large conflagrations remained common in the following century, and when the modern regime conducted building surveys in the 1880s, roughly two thirds of the buildings in the city were still roofed in wood shingles, which Yoshimune's campaign had been

intended to eliminate.[12] The fireproofing program of the 1730's was part of a larger campaign for fiscal retrenchment aimed at returning an overextended ruling elite to solvency. In keeping with the shogunate's conception of economic management as fundamentally a matter of moral discipline, the same edicts that offered loans for fireproofing enjoined samurai to build smaller houses. The policy's aim was thus to stabilize economic and social conditions rather than to build a more secure environment for economic growth. When town elders protested, the regime did not attempt to enforce policy against their wishes. Such accommodation provided a sign of benevolent rule more consistent with Tokugawa political values than slum clearance for the sake of fireproofing would have been.

As a result, the same policy that kept Edo in constant danger also preserved class diversity in the commoners' districts. Obligated to look after their poor tenants and holding little prospect of improving their investment by building more expensive buildings, landlords preserved the status quo and generally sought their profits in commerce rather than rents. This arrangement appeared economically wasteful to political leaders under the new capitalist regime after 1868. In the 1870 s and 1880s, when fireproofing policies were enforced in central Tokyo, the poor were forced out. Tokyo Mayor Matsuda Michiyuki regarded the expulsion frankly as desirable, writing in a letter to the Prime Minister in 1880 that "if those lacking the finances to respond to this system should leave the central districts of their own accord and settle in some other, less appealing district, I would think it only a natural outcome that there would be a turnover of rich and poor and the city would return to peace".[13]

The political objectives we would like to believe lie at the foundation of disaster policies today are democratic ones. Yet the priority on private property and capitalist growth, which in Japan began transforming policy abruptly after the collapse of the Tokugawa regime and then accompanied policies to demonstrate Japan's modernity to the Western powers, created inequities in some cases more severe than under the overtly hierarchical and autocratic Tokugawa regime. The longer history of the Fukushima region and of nuclear power in Japan makes this evident. As sociologist Kainuma Hiroshi has documented in detail, the municipality in Fukushima where the reactors were built in 1970 accepted the power company's proposal, which came with a generous package of economic incentives, in the hopes of salvaging a local economy that had depended on coal mines and had been devastated by the country's shift to imported oil. The region had suffered poverty for generations. In retrospect, we can see that Fukushima residents entered into a Faustian bargain. Yet just like the poor who were pushed out of central Tokyo to make way for fireproof structures, which would in turn show the imperial powers that Japan was building a prosperous modern nation, the residents of Fukushima were not simply embracing nuclear power "of their own accord". The region had been marginalized and exploited throughout Japan's modernization. In what Kainuma rightly describes as a "colonial policy", the politically powerless and vulnerable citizens of Fukushima were then targeted

[12]For statistics on roof materials, see [19].

[13]Matsuda Michiyuki, "Shōshitsu no ato kaoku seigen no gi ni tsuki ukagai," February 16, 1880, reprinted in *Tōkyō shishikō shigai hen 64* (Tōkyō shiyakusho, 1973), 609.

by state and industry to be sacrificed for the advancement of national industrial development [7].

7.3 Conclusion

Despite the gulf that separates the Japan of today from early nineteenth-century Edo, there are some broad commonalities in their responses to risk that suggest what may be human universals. First, both nineteenth-century Edoites and post-Fukushima Japanese act more on the basis of their perceptions of risk and their sense of control over their fates in a hazardous situation than on the basis of the statistical probability of coming to harm. For Edoites, large conflagrations posed a serious threat, but since most of the time they followed predictable patterns and well-trained residents had strategies of escape, they were treated as a spectator sport more often than a cause for panic. Since the meltdown at the Fukushima Daiichi nuclear plant in 2011, residents of a large region around the plant—and even some people overseas—have lived with the daily anxiety that nuclear radiation may be making them sick and shortening their lives. It doesn't matter that scientists' data suggest that this risk is lower than other risks for them; the fact that it is pervasive, invisible, unknowable, and seems to render them powerless prioritizes it psychologically above other risks.

Second, strong bonds among neighbors promote disaster preparedness as well as helping people cope with the effects of a disaster. Edoites practiced a vigilance against the risk of fire that would be impossible in most societies today. It was possible in Edo because of the high level of mutual obligation among commoner residents. Particularly in the countryside, Japanese communities still have a reputation for strong social solidarity. Yet as local, prefectural and national governments have offered different plans for the future of northeast Japan since 2011, deep divisions have emerged within towns in the affected regions over how to rebuild and what precautions to take against future disaster (and particularly what to do about nuclear power plants), showing that the mediation of state institutions and of party politics that characterizes modern democratic society has inevitably attenuated local bonds.

Third, in both the Tokugawa period and in the present day, Japanese have sought to place blame on particular actors after disasters, thereby giving the problem a human face and allowing existing political structures to appear capable of managing it. Rounding up and executing suspected arsonists and punishing householders who failed to extinguish a fire in their block offered a tangible form of control over the fire hazard without fundamentally altering its effects. Much the same may be true of pressure that led to the resignation of Prime Minister Kan Naoto following the 2011 disaster and to some extent even of the litigation against the Tokyo Electric company for its handling of the nuclear plant failure. Although a powerful political movement against nuclear power has emerged in Japan since 2011 (and indeed, former Prime Minister Kan himself is part of it), focus on assigning blame to particular individuals deflects attention away from asking what to do about the hazard-prone and unsus-

tainable society that most Japanese, like their counterparts in other economically advanced countries, have come to take for granted.

On the other hand, there are profound differences between Edo and present-day Japan, and recognizing the political nature of these differences points us to lessons that might be learned—and perhaps applied, at least piecemeal—from the case of flammable Edo. These differences stem ultimately from the basic political difference that unlike the modern state, the Tokugawa state was not seeking to facilitate capital accumulation. First, the Edo approach to building and habitation made what I have called the "technomass" of the city extremely light, which engendered greater flexibility and resilience. Because buildings could be put up in days and partially or wholly dismantled in minutes, and because furnishings were portable, the damage of a large fire in most cases had little long-term effect on the welfare of the population. Edoites accepted the day-to-day risk of having the roof swept from over their heads for long-term security from the specter of death or ruin in a greater disaster. Compared to the post-Fukushima present, one could also say that they were unknowingly also trading the possibility of greater prosperity for security from the unending anxiety created by a technical infrastructure whose threats are boundless, albeit ordinarily improbable.

Second, and also related to the scale and character of the urban infrastructure, the example of Edo shows that given the tools to build and manage their living environment, people can be quite resourceful at managing risk collectively, just as they are at maintaining the social bonds of community. Of course, community exercises in managing risk, such as fire drills, continue today in Japan and elsewhere. What differs is that Edo commoners understood that the entire responsibility to cope with the fire hazard fell to themselves and, at the same time, they possessed the means to cope. There was no superordinate body responsible for controlling fires, but within the limits of living in a fire-prone wooden city, the common people knew how to manage without outside assistance. Engaging people in a limited way in responding to the risks in their environment may be helpful, but the kind of collective coping that was the norm in Edo requires that the environment threatened is one fully controlled by its occupants.

Finally, the contrast between Edo and the present day brings into sharp relief the extent to which modern safety regimes are tied to the value modern societies have placed in private property and capitalist growth. We need not harbor romantic nostalgia for feudalism to recognize that the Tokugawa approach to fire risk had certain social merits as its incidental effects. Since social order was given priority over the preservation of buildings and material goods, there was no slum clearance—indeed, probably no large concentrated slums—and the city avoided economic class segregation. Nor did the protection of private property become a public trust with the potential to distort priorities in coping with hazards. In city building and in disaster management, policy makers and planners today might take hints from both of these features of a society that managed risk very differently than we do, but on its own terms managed it well.

References

1. Beck U (2008) World at Risk. Polity Press, Cambridge
2. Clancey G (2006) Earthquake nation: the cultural politics of Japanese seismicity, 1868–1930. University of California Press, Berkeley
3. Davis M (1999) Ecology of fear: Los Angeles and the imagination of disaster. Metropolitan Books, New York
4. Douglas M, Wildavsky A (1982) Risk and culture: an essay on the selection of technological and environmental dangers. University of California Press, Berkeley
5. Havens T (1978) Valley of darkness: the Japanese people and world war II. W.W. Norton and Company, New York
6. Jannetta AB (1987) Epidemics and mortality in early modern Japan. Princeton University Press, Princeton
7. Kainuma H (2011) *Fukushima ron: genshiryokumura wa naze umareta no ka.* Seidosha, Tokyo
8. Kelly W (1994) Edo and Paris: urban life and the state in the early modern era, chapter Incendiary actions: fires and firefighting in the Shogun's capital and the people's City, pp 310-331. Cornell University Press, Ithaca
9. Kitahara I (1983) Ansei daijishin to minshū: jishin no shakaishi. San'ichi shobō, Tokyo
10. Kitahara I (2013) Jishin no shakaishi: Ansei Daijishin to minshū. Yoshikawa Kōbunkan, Tokyo
11. Kuroki T (1999) Edo no kaji. Doseisha
12. Kyokusai M (1766) Chinka yōjin kuruma
13. Matsunosuke N (1975) *Edo chōnin no kenkyū 4,* chapter Edo no machi nanushi Saitō Gesshin, page 433-434. Yoshikawa kōbunkan
14. Matsunosuke N (1978) *Edo chōnin no kenkyū 5,* chapter Kasai toshi Edo no jittai, p. 28. Yoshikawa kōbunkan
15. Metzler M, Smits G (2010) Economic thought in early modern Japan, chapter Introduction: the autonomy of market activity and the emergence of Keizai thought, pp 13–14. Brill, Leiden
16. Miyata N et al (1995) Namazue: Shinsai to Nihon no bunka. Ribun shuppan, Tokyo
17. Morse ES (1917) Japan day by day, vol 1. Houghton Mifflin Co, Boston
18. Nishida Y (2006) Kōshō: Edo no kasai higai wa sukunakatta no ka. Jūtaku shinpōsha, Tokyo
19. Ogi S (1979) Tōkei shomin seikatsu no kenkyū. Nihon hōsō shuppan kyōkai, Tokyo
20. Smits G (2006) Shaking up Japan: edo society and the 1855 catfish picture prints. J. Soc. Hist. 39(4):1045–1078

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

Chapter 8
Conclusion

Corinne Bieder

Pretending to come to definite conclusions on uncertainty or uncertainties and how to live with it or them as safely as possible would have been pretentious. After all, uncertainty embodies to some extent the unknown. However, if the concept itself was only introduced as a central concept to risk management or safety in recent years, mankind has long lived and is still living with numerous uncertainties. The question is then whether it is doing so safely enough. Are there successful or satisfactory experiences of organizations dealing with uncertainty? In such cases, are they consciously and deliberately addressing uncertainty? Can their approaches be described, analyzed and extended to all hazardous activities?

8.1 Uncertainty: A Multi-faceted Notion

As aforementioned, uncertainty is increasingly frequently used as a concept to move forward in the risk or safety management domains (see ISO 31000, definition of risk management). Yet, trying to dig into what it really refers to is complex. Confronting perspectives on uncertainty make it obvious that the concept, as vague as it is today, requires some conceptual clarification and refinement. Indeed, it turns out that it may in reality refer to a variety of aspects. Furthermore, it is far from being an absolute notion.

When used by psychologists, uncertainty is likely to refer to the incomplete information that an operator has on the situation (s)he has to cope with.

For the residents of Edo back in the 17th century, it was the time at which a fire would occur as well as the wind direction which, to some extent (see Chap. 7), corresponds to their "operator's" incomplete information on the situation.

C. Bieder (✉)
Ecole Nationale de l'Aviation Civile (ENAC), Toulouse, France
e-mail: corinne.bieder@enac.fr

© The Author(s) 2017

G. Motet and C. Bieder (eds.), *The Illusion of Risk Control*,
Safety Management, DOI 10.1007/978-3-319-32939-0_8

From a risk management point of view, uncertainty can be related either to the absence of information on the consequences of an event and their severity, or to the absence of numbers characterizing its likelihood. The CERN example is a striking illustration of the uncertainty of quantification (see Chap. 6).

From a broader perspective, uncertainty also covers unknown-unknowns, things that no one has ever anticipated or even imagined. Here again, it could be the combination or simultaneity of known individual phenomena or the phenomena themselves that no one has ever imagined, or an evolution at a time horizon that no one has ever considered. Human-induced climate change has only recently appeared on radar screens when anticipating the impact of technological and societal evolution.

While the word "uncertainty" has emerged as a new magic wand in the safety community, the concept itself has not been fully clarified in theory and its numerous interpretations, especially the dimensions considered, are most of the time left implicit.

8.2 Uncertainty: A Relative Notion

The multi-faceted nature of uncertainty is unfortunately (or maybe fortunately) not the only subtlety of the concept. Another one that seems often overlooked is its relativity. Indeed, anticipating hazards is a creative process more than a universal, well established approach. As such, it involves a number of aspects including experience, imagination and others. Some people or organizations may anticipate far more hazards and accident scenarios than others. How this kind of uncertainty can be accounted for is still an open question.

Ultimately, things may be a surprise for some, but not for others. Whistle blowers think they are certain of things that others deny. Likewise, a fire in a remote area could be completely unexpected by all, except the pyromaniac who lit it. In other words, we need to ask who is uncertain. When we commonly say after an event "no one could have imagined it", what does this mean in reality? No one? No one we know of? No one who could have said or done something or who was in a position to say or do something?

In addition to varying from one person or organization to another, uncertainty also varies from one moment in time to another. As knowledge and experience grow, uncertainty evolves. It is not necessarily reduced though, conversely to what one could intuitively believe. Indeed, new knowledge may open up new areas of unknowns. When elementary particles were discovered, the whole area of quantum physics was unsuspected.

In other words, uncertainty turns out to be a relative more than an absolute notion. Relative to persons, relative to time... at least.

From a more philosophical point of view, uncertainty and risk share the same difficulty. They relate to the future, but we attempt to characterize and measure them with current knowledge. From an epistemological perspective, we can never accumulate

enough knowledge to achieve it (see Chap. 2). While the assessor's uncertainty can be reduced, the assessment itself will always be debatable.

The existence of debate and controversy is sometimes used to characterize uncertainty as the embodiment of the relativity of the notion. This may lead to considering that there is no uncertainty related to an idea/approach/action if everyone agrees, even if this idea/approach/action turns out to be erroneous or uncertain. Furthermore, such a characterization of uncertainty doesn't include unknown-unknowns in its scope.

Eventually, Uncertainty with a capital U becomes meaningless, making the need for some further conceptual work even more obvious. However, can we wait for additional conceptual development to be achieved (and what could be expected from it?) before starting to deal with uncertainty? Despite the current conceptual shortcomings, a number of approaches have been and are still being developed to live with uncertainty as safely as possible.

8.3 Reconciling Risk Management and Social Science Approaches to Uncertainty

If early risk management methods assumed that it was possible to reduce uncertainty to a point where it would no longer exist, recent reflections (and facts!) on uncertainty seem to converge at least on one aspect: we cannot and shouldn't only fight uncertainty.

Up to now, the approaches developed based on classical risk management methods and the ones derived from the social sciences used to be opposed to one another for their fundamental conceptual differences and also, if not more, for their difference in underlying disciplines.

Classical risk management methods, derived from engineering approaches, used to assume that uncertainty can, if not be reduced to zero, at least be described through a limited number of scenarios that can be anticipated and their likelihood calculated. With progress in computing capacity, the number of scenarios envisaged has increased over the years, further encouraging quantitative approaches of uncertainty and reinforcing the "illusion of control".

However, recent risk management method developments introduce an additional dimension or parameter to the likelihood and severity of consequences of risk, meant to reflect the uncertainties embedded in these dimensions without really characterizing them although trying to measure them (see Chap. 3). If uncertainty is getting closer to the world of the unknown, it is still expressed using equations.

Taking a closer look at it allows us to characterize the kind of uncertainty accounted for by these approaches. It seems to refer to the domain of the knowable or the known-unknowns, at least from the perspective of those performing the risk analysis at the time they are performing it.

Approaches deriving from social sciences have never attempted to describe or characterize uncertainty as such. Yet, the approaches deriving from these disciplines tackle it in different ways.

Some approaches such as organizational learning try to extend the world of the known to the whole organization rather to leave it only to those who experienced a situation. The experience accumulated, be it directly or indirectly through discussions with peers or other practices, allows for pushing back the frontiers of some kind of uncertainty, the knowns to some—unknowns to others at a given point in time.

The resilience school of thought suggests that it is more about how to live with unknowns to the organization when faced with them. However, they are not pretending to reduce the world of the unknown-unknowns.

Eventually, pushing the discussion further between risk management and social sciences representatives leads to a much more optimistic conclusion than the initial somehow sterile monologues. Indeed, while both are referring to uncertainty, their scopes are different but interestingly enough complementary.

The theoretical scope of risk management is limited to the world of the known and knowable to the risk managers and the experts they are relying on at the time when they perform the risk analysis. Therefore, if unknown-unknowns to anyone at the time considered are not satisfactorily addressed by risk management approaches, it is essentially because they are outside of their theoretical scope.

However, unknown-unknowns do exist and organizations have to be prepared to cope with them. Unlike the turkey that gets used to be comfortably fed for 364 days and is caught by surprise on Thanksgiving day [2], resilience engineering approaches aim at making organizations resilient to these unexpected or even unexpectable situations (see Chap. 4).

Everyone then agrees that it cannot be done by means of an improved risk analysis method.

In summary, both risk analysis and social sciences based approaches such as resilience, pragmatist theory or collective mindfulness, or to refer to another dichotomy of approaches, Safety 1 and Safety 2, are doing the job they are meant to do with respect to uncertainty. Of course they can be improved in their own field. But the mismatches lie in the kind of uncertainty they are respectively contemplating. Any refinement, as sophisticated as it might be, won't change the scope of reference. What could be improved is their use through better characterizing what to use them for and what can be expected from them. In this perspective, risk analysis becomes a reasonable argument more than an objective representation of the truth, useful for dialectical debate on safety.

8.4 Should We Fear Uncertainty?

Whatever the approach to uncertainty, thus whatever the type of uncertainty, the very fact of associating it with the concept of risk has induced a certain fear of uncertainty. Yet, strangely enough, humans are used to living with uncertainty, at least some kind

of uncertainty. It is the case for example of the incomplete information we have in situations we are dealing with. Most people make their way safely through life despite a number of major uncertainties and unexpected situations. A way forward could be to build on this existing ability at a broader scale.

Why are we so afraid of uncertainty? Has it always been the case? The residents of Edo chose to stay there and live there. They didn't flee from the city and its repeated fires. Closer to us, there is no major reluctance from the general public to use mobile phones despite today's uncertainty on the electromagnetic hazards of these devices. There are numerous examples of known-unknowns that we all live with, sometimes willingly. Why is it so?

It seems that for the accepted risks or uncertainty the associated benefits are clearly perceived as such. How many people feel lost when they forget their mobile phones or when the network is down? Who would envisage spending a week on a boat to cross the Atlantic one way (except for cruise addicts)? However, is this risk/benefit balance sufficient to make a risk or uncertainty widely accepted? In fact, the alchemy seems to be more subtle. Even when uncertainty is not perceived only negatively, there appears to be some conditions to make it possible to live with it. Even though most people are not ready to significantly reduce their electrical power consumption in countries where nuclear power is used, there are strong opponents to the nuclear industry. What else is then needed beyond perceived benefits to make uncertainty more acceptable to most people?

The Edo case described in Chap. 7 highlights at least two complementary major conditions:

- giving people some control over their fates;
- the hierarchy of values is respected in the way risk is managed.

If we consider how the general public is currently involved in risk management or can feel in control, it is mainly through intermediaries, be they Regulators or other more or less formal organized bodies. Feeling in control and involved thus necessarily relies on a key element: Trust.

More generally, at all levels the question of trust appears as a central one when it comes to living with uncertainty, especially when the magnitude of potential consequences is high. Trust in experts, trust in decision-makers, trust in Authorities and all the ones who may have different interests and/or a different perimeter of uncertainty.

8.5 Living with Uncertainty: Beyond Organizations and Regulators...

If innovative hazardous activities tend to learn from classical ones such as nuclear, aviation..., they can also help us further reflect on what is needed to build trust and ultimately live with uncertainty in other hazardous industries. The example of gene therapy research, acknowledged as being a hazardous activity with a high level of

uncertainty, brings a new perspective on this question. More precisely, it broadens the scope of the question to its societal and ethical dimensions.

Indeed, it seems today that most research on new risk management approaches based on uncertainty focuses largely on organizations themselves (e.g. resilience) or tackle regulation regimes. However, hazardous industries operate in a much broader framework. Modern society models have created a pathological relation to risk, the illusion of control. They changed mankind's relation to risk.

Acknowledging that risk/uncertainty management is embedded in a political context/process—exerting a significant influence—led gene therapy research to getting back to fundamental questions: the questions of moral responsibility and societal benefits. Eventually, it also led to evolving towards a morally responsible approach to risk/uncertainty management [1].

If the questions of moral responsibility and societal benefits were certainly raised at the advent of now well-established hazardous technologies, acknowledging uncertainty as part of hazardous industries' operations may require revisiting these questions and integrating them into a broad dynamic societal risk management framework.

References

1. Baram MS (2004) Managing the risks of testing new gene therapies on humans. How to manage experience sharing: from organizational surprises to organizational knowledge. Elsevier, Amsterdam
2. Taleb NN (2007) The black swan: the impact of the highly improbable. Random House, New York

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.