EFFECTIVENESS OF RISK WORKSHOPS: AN AGENT-BASED SIMULATION STUDY

Universität Hamburg

Fakultät für Wirtschafts- und Sozialwissenschaften

Dissertation

Zur Erlangung der Würde des Doktors der Wirtschafts- und Sozialwissenschaften

(gemäß der Promotionsordnung vom 18. Januar 2017)

vorgelegt von

Clemens Harten

aus Hamburg

Hamburg, 2025

Vorsitzender:	Prof. Sebastian Späth
Erstgutachter:	Prof. Matthias Meyer
Zweitgutachter:	Prof. Dominik Vogel

Datum der Disputation: 03. Mai 2024

Danksagung

Die vorliegende Dissertation wäre nicht möglich gewesen ohne die Unterstützung vieler Menschen. An dieser Stelle möchte ich ihnen meinen Dank aussprechen.

Mein besonderer Dank gilt meinem Doktorvater, Prof. Dr. Matthias Meyer, der mir mit seiner großen fachlichen Expertise und seinem Vertrauen in meine Arbeit eine wichtige Unterstützung war. Seine Anregungen und sein Einsatz haben entscheidend zur Qualität der Arbeit beigetragen und diese überhaupt erst möglich gemacht. Mein Interesse am wissenschaftlichen Arbeiten geht auf den Kontakt zu ihm und seinem Institut zurück.

Gleichermaßen danke ich Prof. Dr. Lucia Bellora-Bienengräber, die diese Arbeit unermüdlich gefördert hat. Ihre Perspektive hat mich und diese Arbeit sehr bereichert. Beiden, Prof. Meyer und Prof. Bellora-Bienengräber, danke ich auch für die äußerst engagierte Zusammenarbeit bei der Erstellung der Paper, die im Rahmen dieser Dissertation entstanden sind.

Weiterhin danke ich Prof. Dr. Dominik Vogel für die Erstellung des Zweitgutachtens und Prof. Dr. Sebastian Späth für die Übernahme des Vorsitzes des Verfahrens.

Meinen Kolleginnen und Kollegen am Institute of Management Accounting and Simulation der TU Hamburg danke ich für die vielen konstruktiven Diskussionen, die meine Arbeit bereichert haben. Ganz besonders danke ich Prof. Dr. Iris Lorscheid, die mir eine große Hilfe bei der Findung meines Themas war und von der ich viele wichtige Impulse bekommen habe. Prof. Dr. Sina Völtzer, Dr. Kai G. Mertens, Dr. Jonas Hauke und Sebastian Achter danke ich für ihre Ideen und ihren Zuspruch.

Mein Dank gilt auch den vielen Menschen mit denen ich auf Konferenzen und Workshops, per Mail oder auf Universitätsfluren über meine Arbeit sprechen konnte. Ich danke der ERRN- und der ESSA-Community, der Risk Governance Research Group der Universität Siegen, Dr. Jan Spitzner und meinen Kolleginnen und Kollegen am HARCIS.

Abschließend danke ich meiner Frau Julia für ihre bedingungslose Unterstützung, ihr Vertrauen in mich und ihren ausdauernden Rückhalt. Ihr und unseren wunderbaren Kindern Charlotte und Laura widme ich diese Arbeit.

Hamburg, im Juli 2025

Clemens Harten

Table of Contents

page

List of Abbreviations

ABM	Agent-based modelling
BN	Bayesian network
cf.	compare
CoMSES	Computational Modeling in the Social and Ecological Sciences
COSO	Committee of Sponsoring Organizations
c.p.	ceteris paribus
CRO	chief risk officer
CSN	constraint satisfaction network
DAG	directed acyclic graph
ECHO	explanatory coherence by harmany optimization
ERM	enterprise risk management
e.g.	exempli gratia
GPL	General Public License
i.e.	id est
ISO	International Organization for Standardization
ODD+D	Overview, Design Concepts and Details + Decision
SEC	Securities and Exchange Commission
VaR	value at risk

List of Figures

page

List of Tables

page

1 Introduction

Risk management has become a crucial aspect of modern organizations (Aven, 2012; Power, 2004). Managing risks effectively enables organizations to navigate uncertainty, capitalize on opportunities, and protect their financial stability and reputation. As organizations face a multitude of potential threats, ranging from financial and operational to legal and reputational, adopting a proactive risk management strategy is essential to ensure their long-term success and sustainability (Hopkin and Thompson, 2022).

To aid organizations in their risk management efforts, numerous techniques, recommendations, and standards have been developed. These resources provide businesses with valuable guidance and best practices for identifying, assessing, and mitigating risks associated with their operations. By leveraging these tools, organizations can create tailored risk management frameworks that address their specific needs and challenges, thus enhancing their decision-making processes, optimizing resource allocation, and bolstering their resilience to external shocks (Hunziker, 2019; Kewell and Linsley, 2017; Ostrom and Wilhelmsen, 2019).

Organizations aim to employ best practices for risk management to ensure their success in a competitive landscape. By adhering to established standards and recommendations, businesses can demonstrate a commitment to responsible management and strengthen stakeholder confidence. Furthermore, implementing effective risk management practices can help organizations maintain regulatory compliance, minimize legal liabilities, and foster a culture of risk awareness throughout the organization.

As organizations rely on best practices and recommendations for risk management, it is crucial that sound, up-to-date advice is available to facilitate effective decision-making (Bromiley *et al.*, 2014). Reputable sources, such as industry experts, regulatory bodies, and professional associations, play a vital role in developing and updating guidelines to ensure that organizations have access to the latest knowledge and best practices (e.g. Aven and Renn, 2010; COSO, 2004; ISO, 2009; Rausand, 2011). Academic research can contribute by testing the actual effectiveness of advice and best practices in risk management. By, for example, conducting empirical studies or analyzing real-world scenarios, researchers can validate the effectiveness of existing guidelines, identify potential gaps, and contribute to the development of more robust and comprehensive risk management advice (e.g. Lucas, 2001). This, in turn, enables organizations to make informed decisions regarding the risks they face.

Despite the importance of risk management for organizations, there is still a notable lack of knowledge regarding the actual effectiveness of various risk management practices and how they should be implemented (Aven, 2012; Bromiley *et al.*, 2014; Fraser *et al.*, 2009; Schmit and Roth, 1990).[1] This lack of knowledge can hinder organizations in their quest for effective risk management, as they may struggle to determine which practices are most suitable for their unique context and needs (Fraser *et al.*, 2009). Consequently, it is crucial to bridge this knowledge gap. Academic research can contribute to the development of more effective risk management practices by evaluating the proposed practices.

Researching the effectiveness of risk management practices, however, presents many challenges. One of the primary obstacles is the inherent difficulty in differentiating between effective and ineffective practices (McGrew and Bilotta, 2000). The effectiveness of the entirety of the risk management of an organization can be evaluated through a number of measures, each with its own strengths and limitations. For example, one can scrutinize the alignment between the risk management strategies employed and the organization's overall objectives (Beasley and Frigo, 2007; Sobel and Reding, 2004). An effective risk management practice should demonstrably contribute to the achievement of these objectives. Additionally, the frequency and severity of unexpected events or losses can be used as indicators of risk management effectiveness (Zwikael and Ahn, 2011). A lower incidence rate may signify a more robust risk management system. Furthermore, the responsiveness to and recovery time from adverse events can serve as metrics for risk management effectiveness. The quicker an organization can return to its normal operations after a disruption, the more effective its risk management practices may be considered (Kumar and Park, 2019; Sodhi and Tang, 2009). Finally, an organization's level of compliance with industry standards and regulations can also provide insights into the effectiveness of its risk management practices (Lundqvist, 2014; Tamimi, 2021).

[1] Bromiley *et al.* (2014, p. 273) find that "[t]o contribute to the ongoing ERM discussion, management scholars need to take a more prescriptive stance and pay more attention to the effectiveness of different practices and activities." Fraser *et al.* (2009, p. 399) see a "critical need for more detailed 'real-world' applications on ERM". Furthermore, they provide the assessment of a practitioner (Fraser *et al.*, 2009, p. 399) that "[t]he impact of corporate culture on ERM implementation and practices is not well addressed in the literature."

However, since risk management is often deeply integrated into an organization's operations, isolating the effect of *individual practices* for scrutiny can be a difficult task. This is further complicated by the fact that the impacts of risk management decisions often materialize over long periods, making it difficult to establish causal links (Schmit and Roth, 1990). Consequently, it becomes difficult to determine whether a successful outcome is a product of effective risk management practice, luck, or other organizational processes.

In previous research, especially laboratory experiments have been used to investigate individual risk management practices. For example, Lee *et al.* (2019) investigate the impact of mental construal on risk management activities in a web-based experiment with IT project managers. Keil *et al.* (2000) use a laboratory experiment to measure how risk-related information affects the risk assessment of software development projects. O'Donnell and Prather-Kinsey (2010) compare the risk assessment of auditors in standardized tasks to identify the impact of the auditors' nationality. Beyond the investigation of risk management practice, laboratory experiments have been conducted extensively to investigate individuals' risk perception and risk attitude and their associated behavior (e.g., Arbis *et al.*, 2016; Bajtelsmit *et al.*, 2015; Mear and Firth, 1988; Weber and Hsee, 1998).

While the previously mentioned experiments focus on the decision-making of individuals, some studies have also performed laboratory experiments with groups that collaborate in risk management-related tasks. For example, Valacich *et al.* (2009) compare the decision-making of groups depending on the risk preference of the firm they work in, the information distribution within the team, and the mode of communication (computer-mediated or face-to-face). Vriezekolk *et al.* (2015) assess the reliability of a specific risk assessment method (the RASTER method) by asking teams of participants in a laboratory experiment to collaboratively assess risks.

Laboratory experiments allow researchers to control the environment and context in which participants make decisions. As much of risk management practice can be described as information processing and decision-making (Fenton and Neil, 2019), a main benefit of laboratory experiments to research risk management practices is the possibility to control the information available to the experiments' participants and to record the decisions made by the participants based on the information provided to them in a structured format (e.g., the experiment can provide specific information regarding a risk and ask for a risk assessment within a fixed framework).

While laboratory experiments have been successfully used to investigate risk management practices, they share some limitations that are difficult to overcome in laboratory experiments with human participants. First, experiments that involve human participants are limited with respect to the number of participants and experimental treatments that can be included in a study, due to the financial costs associated with the method (Friedman and Sunder, 1994, p. 5).[2] This makes it difficult to detect small effects that require a high number of repetitions to become evident. And second, while the information provided to the participants can be controlled and the participants can be surveyed to reveal their thoughts, it is still difficult to know what the participants are thinking at each point of the experiment, e.g., at the moment they change their mind regarding a risk assessment (Nisbett and Wilson, 1977).

One possible method to evaluate the effectiveness of risk management practices that is not limited in this regard is the use of simulation experiments. Simulation experiments provide a controlled environment in which the effects of different interventions can be closely observed and evaluated (Axelrod, 1997). This method provides a way to understand and quantify the effectiveness of different risk management practices by systematically applying them in a controlled environment with a potentially unlimited number of repetitions.

So far, there is a lack of research that capitalizes on the benefits of simulation studies to further expand our understanding of how to effectively manage risks. One of the core benefits of simulation experiments is the possibility to control complex systems with many potential variables and interactions (Axelrod, 1997), like a face-to-face group discussion. Because simulation experiments do not experiment with the subject itself but a model of the subject (Gilbert and Troitzsch, 2005, p. 15), they provide full control of the experiment, which makes them well suited especially for complex systems like organizations (Harrison *et al.*, 2007). Simulation experiments can be used for theory development (Davis *et al.*, 2007), to reveal relationships between variables (prediction), to test possible mechanisms that lead to observed relationships (explanation), or to suggest improvements for organizational procedures (prescription) (Harrison *et al.*, 2007).

[2] Sometimes, laboratory experiments are highlighted to have low costs, compared with other methods (Falk and Heckman, 2009). We compare costs of conducting laboratory experiments to those of simulation experiments, where costs of raising the number of repetitions are neglectable, compared to laboratory experiments.

In the context of researching risk management practices, simulation experiments are especially interesting because the cognition of the involved participants (e.g., how they weigh specific information regarding a risk) plays a crucial role in the risk management process. In simulation experiments, a cognitive architecture is implemented to represent the actual cognition. The cognitive architecture allows a detailed understanding of how information is processed and how it influences the decision-making of the (simulated) individual (Sun and Naveh, 2004).

Conducting a simulation experiment requires a well-specified setting in order to build a model that is sufficiently similar to reality (Gilbert and Troitzsch, 2005, p. 15). For the purposes of this study, we choose a setting that is a typical component of risk management in organizations: The assessment of risks by a group of experts during a risk workshop.

Risk workshops are a method to gather perspectives and information from multiple stakeholders to inform a risk assessment (COSO, 2017). They are commonly used in practice and play a key role in the risk management process (Quail, 2011). However, while some guidance has been published on how to facilitate a risk workshop (Hunziker, 2019; Quail, 2011), little research is available that evaluates the effectiveness and impact of the advice provided.

Simulation experiments are potentially well-suited to investigate risk workshops, as they allow us to monitor and control in detail the characteristics of participants, their cognition and behavior, as well as their interaction with each other. Modeling a risk workshop requires choices regarding the implementation of two critical systems: The individual cognition of the participants and the interaction of the participants with each other:

- **Model of interaction**: The model of interaction concerns the exchange between the participants of a risk workshop. In a typical face-to-face workshop, participants will talk to each other regarding a particular risk and share their individual risk assessments. A well-established method to model the interaction of (human) actors for simulation experiments is agent-based modelling (Gilbert and Troitzsch, 2005; Railsback and Grimm, 2011). It has already been applied to simulate group work (e.g., Lorscheid and Meyer, 2021; Son and Rojas, 2011), but not risk workshops specifically.
- **Model of cognition**: Each agent in an agent-based model needs at least some rules that determine its behavior and that model its cognition. The complexity of the

model of cognition depends on the requirements of the simulation experiment (Gilbert, 2005). A cognitive architecture is needed when both the social and the cognitive levels are important to the modeled setting (Gilbert, 2005). Choosing a cognitive architecture for agents in an agent-based model can be challenging. However, we can build upon work from the field of risk management that is concerned with the modeling of risks (e.g., Fenton and Neil, 2019) to build a model of how agents process information about risks.

Using the model we build from these two components – the model of interaction within a risk workshop and the model of the participants' cognition –we aim to answer the following research question: *How to facilitate an effective risk workshop*[3]? The model allows us to perform simulation experiments on risk workshops where we measure the impact of group characteristics, individual behavior, decision-making rules, and choices of the workshop facilitator on the effectiveness of the workshop.

Another factor that impacts risk management practice, but which is not covered explicitly by the first simulation study, is culture (Ring *et al.*, 2016). The impact of culture on risk management practice has been identified as a question in need of further research (Bromiley *et al.*, 2014; Fraser *et al.*, 2009). While 'culture' can refer to different concepts in the context of risk management (Ashby *et al.*, 2012; Bromiley *et al.*, 2014), we focus on the calculative culture (Mikes, 2009; Power, 2007) of the organization that performs the workshop, which can be modeled by adapting the cognitive architecture of the workshop participants. Therefore, we perform two simulation experiments using two distinct models that correspond to the two calculative cultures that Mikes (2009) calls 'quantitative skepticism' and 'quantitative enthusiasm.' This allows us to answer a second research question: *How should a risk workshop account for the predominant calculative culture in order to be effective?*

Chapter 4 describes the results of the simulation experiment under quantitative enthusiasm. Chapter 5 describes the results of the simulation experiment under quantitative skepticism and compares the results of both experiments.

[3] We understand effectiveness of a risk workshop as the workshop's ability to reach a correct risk assessment within limited time.

In the following Chapter 2, I provide the theoretical background of this study, which builds upon previous research on risk and risk management, before I discuss the methodology used for the study in Chapter 3. Chapter 4 and Chapter 5 will present the results of the two simulation studies that were conducted for this thesis.[4]

[4] Additionally, Appendix 1 provides a glossary of important terms used throughout the thesis. Appendix 6 contains a brief summary of the thesis in German and English. Appendix 7 provides a list of publications that are associated with the thesis project.

2 Theoretical Background

2.1 Risks and uncertainty

Risks are omnipresent in modern society, so much so that sociologist Ulrich Beck has described our contemporary society as a "risk society," the successor of the "industrial society" (Beck, 1992). Beck (1992) argues that modern society's production of wealth comes with a production of risks. These risks need to be distributed within society, and discussion and negotiation related to the distribution of risks have become a fundamental aspect of public discourse, which results in a high risk awareness for both individuals and institutions.

Giddens (1999) highlights that the concept of risk is a relatively recent one. While people throughout human history have always faced dangers and hazards, risk is more than that. It is an inherently future-oriented concept and links future hazards to current decisions that potentially affect them (Giddens, 1999). Thus, uncertainty about the future and the possibility of influencing the future are cornerstones of the idea of risk. For example, a gambler must decide whether to take a bet without knowing the outcome. However, the gambler might reason about the probability of specific outcomes and can use that to assess the risk associated with the bet.

Luhmann (1993) describes that technological progress transforms hazards into risks by providing means to avoid the hazard: The existence of umbrellas makes 'getting wet' a risk one takes by not carrying an umbrella (Luhmann, 1993).

While there is no universally accepted definition of the term 'risk' (Aven, 2012; Luhmann, 1993), and even the origin of the modern meaning is disputed (Beck and Kewell, 2014), there is a common understanding that risks are related to decisions that impact future events. This directly leads to the question of how to make good decisions in the presence of risks. In the following, I will discuss how risks have been conceptualized to make them accessible to a systematic approach towards managing risks.

2.1.1 Types of risks

Several classifications of risks have been proposed over time, focusing on different aspects of risks. In the following, I discuss the classification of risks regarding the knowledge associated with them, the threats they are linked to, and their origin. Understanding the different types of risks is necessary to build a model of a risk that matches the risks an organization actually faces.

One possibility is to classify risks based on knowledge and understanding of the underlying problem. For example, the Society for Risk Analysis (Aven *et al.*, 2023) distinguishes between simple and complex, uncertain and ambiguous risk problems (Aven and Renn, 2010).

The distinction between simple and complex risk problems concerns the difficulty of predicting the outcome of events, which is based on the quality of understanding the underlying mechanisms governing the relationship between actions or events and their results. Aven and Renn (2020) give smoking as an example of a simple risk problem (it is well-understood that smoking will lead to lung cancer with a certain probability). Their example for complex risk problems is critical infrastructure systems (for example, events in an electrical grid might have surprising and unforeseen consequences because of unknown or poorly understood interdependencies).

Uncertainty of risk problems refers to the difficulty associated with correctly predicting both the occurrence and the outcome of events or actions (Aven and Renn, 2020). For example, while smoking is a simple risk problem, it might be highly uncertain whether a specific individual will develop lung cancer due to smoking (uncertainty of occurrence). Likewise, in the case of a malfunctioning component of an electrical grid, it might be uncertain how the malfunction will influence the remainder of the grid (uncertainty of consequences).

Uncertainties are often further classified as epistemic or aleatory uncertainty, depending on the source of uncertainty. While aleatory uncertainty is based on true randomness (the uncertainty can be described statistically but not reduced), epistemic uncertainty is based on a lack of knowledge (the uncertainty can be reduced by improving the understanding of the problem) (Hora, 1996).

If a risk problem is ambiguous, there is more than one possible view of the available information. Aven and Renn (2010) distinguish interpretative ambiguity, concerning the "relevance, meaning and implications" of information, and normative ambiguity, concerning the "values to be protected and the priorities to be made." For example, while a company might know that customers oppose raised prices, it can still be unclear if that translates into customers changing their purchase decisions (interpretative ambiguity) or if alienating the most price-sensitive customers is actually harmful (normative ambiguity).

The classification of risk problems mentioned above focuses on the understanding of the relevant information. Another approach to classifying risks is to distinguish the type of threat (or opportunity) associated with the risk. Hopkin and Thompson (2022) distinguish compliance risks, hazard risks, control risks, and opportunity risks. Compliance risks are derived from the laws a company has to adhere to and the associated penalties. Hazard risks are strictly negative events an organization is exposed to due to its activities, e.g., health risks to employees or consumers. Control risks arise from uncertainty about future events. Their outcome is hard to quantify, predict and control. Finally, opportunity risks are associated with potentially positive outcomes of taking risks, e.g., entering new markets. Hopkin and Thompson (2022) highlight that each of these risk types is usually met by organizations with different actions, as compliance risks need to be minimized, hazard risks need to be mitigated, and control risks managed. Opportunity risks, however, should be embraced (Hopkin and Thompson, 2022).

A third dimension of risks is their origin from the perspective of an organization. For example, the PMBOK (Project Management Body of Knowledge) distinguishes "technical, quality and performance" risks related to the product, project management risks, and organizational risks focused on processes within the organization and external risks that originate from outside the organization, like lawsuits or extreme weather events (Pritchard, 2015).

The diversity of risk classifications highlights the importance for practitioners to think about risks in a structured manner. The classifications help with identifying relevant risks (e.g., by actively searching for exposure to compliance risks), identifying the measures needed to improve the organization's understanding of risks (e.g., by differentiating uncertainty and ambiguity), and guiding how to best address the risk (e.g., by improving internal processes).

2.1.2 Risk, uncertainty, and knowledge

There have also been attempts to identify the underlying structure of any individual risk. Aven (2019, p. 58) describes risk as a triplet of a possible future event (A), the consequences of that event (C), and the uncertainty regarding the consequences (U). For example, a new product introduced into the market (A) might prove unsafe for consumers, harming the producer's reputation and requiring a product recall (C). The probability of releasing a hazardous product is unknown (U). Thus, the risk is defined as (A, C, U). Figure 1 provides an overview of concepts and their relationships.

The consequences of an event (C) and the uncertainty (U) must be specified to assess a risk. The specification is built upon knowledge about the risk under assessment (K). The specification of the consequences happens by defining concrete measures of interest (C'). The uncertainty is expressed by a judgment of both the likelihood of C' and the reliability of the knowledge K, codified as Q (Aven, 2019, p. 61). Thus, the full description of the risk is the triplet (C', Q, K). In the previous example, one specific consequence would be the necessity of a recall (C'), the probability of which (Q) could be quantified by consulting the engineers responsible for product development, safety experts, or certification agencies (K).

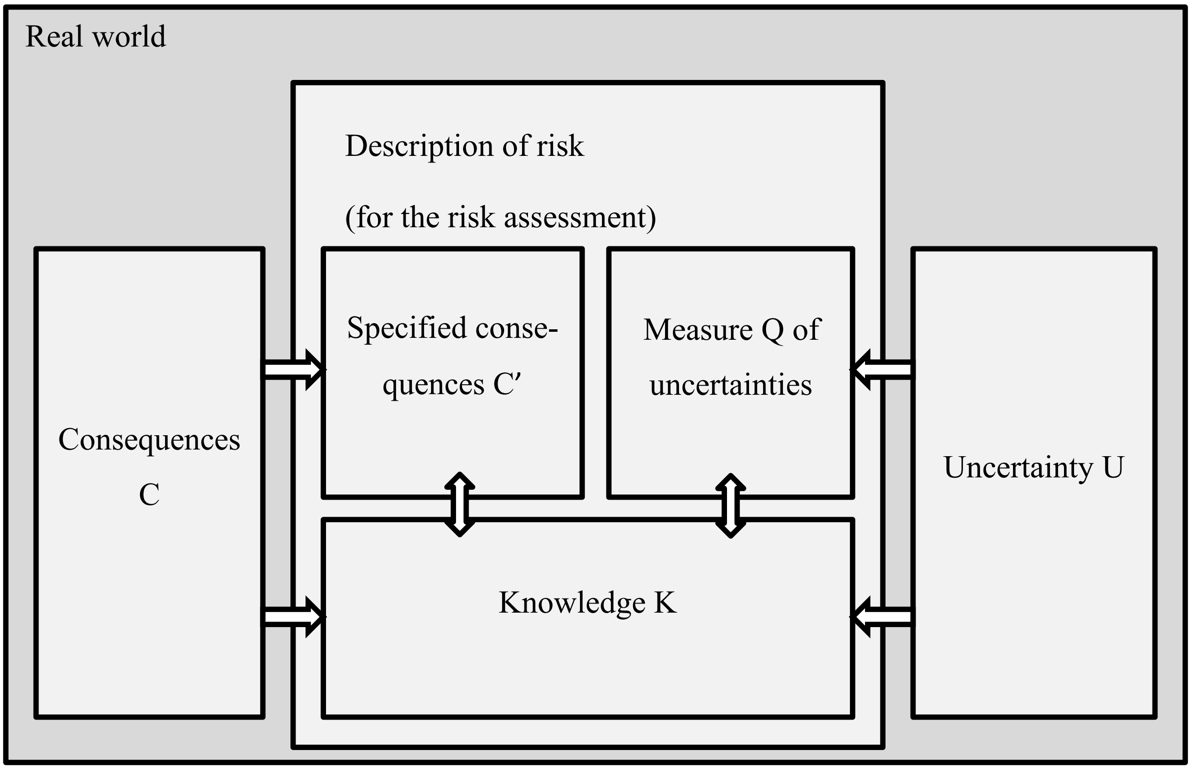


Figure 1 Relationship of risk, uncertainty, knowledge, and consequences.[5]

This framework leads to another interest regarding risks besides understanding their context: the interest in measuring them. This is necessary to make well-informed decisions regarding potential actions related to the risk.

2.1.3 Measuring risks

Many risk management activities require that the risks are measured and ranked according to their severity. Several measures of risk severity are in use, depending on the type of risks under investigation. Most risk severity measures measure either the likelihood or impact of risks. Regarding the measure of likelihood, the COSO framework (2017) distinguishes qualitative descriptions (e.g., "a remote possibility"), quantitative descriptions (e.g., a probability of 80%), and frequency descriptions (e.g., "once a year"). For example, an organization might want to reduce the frequency of workplace accidents and lower the probability of a strategic risk.

Regarding the measure of risk impacts, the COSO framework (2017) gives the example of distinguishing financial impact (e.g., the expected growth is not met) from an impact on operations (e.g., not enough staff is hired to operate at full capacity).

[5] Figure adapted from Aven (2019).

Often, the severity of risks is evaluated using several measures, e.g., one for the likelihood and one for the impact of the risk. These measures allow to judge risks on their own and relative to other risks, as well as providing a measure for potential actions taken by the organization (e.g., reducing the frequency of accidents by investing in appropriate training) (COSO framework, 2017).

Some risks (foremost financial and operational risks) can be quantified regarding an associated financial gain or loss. These risks are often measured regarding their value at risk (VaR), that is, the "maximum expected loss, given some time horizon and within a given confidence interval" (Olson and Wu, 2020, p. 79). The VaR combines a risk's impact and likelihood into one measure that allows the direct comparison of portfolios of (financial) risks (Olson and Wu, 2020).

These risk measures define the output of the risk assessment and are therefore important for measuring risk management practice effectiveness: An effective risk workshop, for example, is able to determine the severity of risks correctly.

2.2 Enterprise risk management

Organizations constantly face risks, with consequences ranging from minor to catastrophic. In order to react appropriately to these risks, standard practices have been developed over the past decades (Bromiley *et al.*, 2014; COSO, 2004; Hunziker, 2019). This provides a certain amount of consistency in the risk work (Power, 2016) performed by the organizations. In the following, the common processes and tools in enterprise risk management will be presented as they inform the simulation models created for this thesis.

2.2.1 The risk management process

The task of understanding and managing risks is central to a successful organization. The field of enterprise risk management has emerged to investigate and propagate best practices for risk management in organizations. Several standards and frameworks have been developed in the past decades to identify the relevant risks, correctly assess them, and derive appropriate actions, most prominently the COSO framework (2004) and ISO 31000 (2009).

Both frameworks provide similar concepts for the risk management process. The following provides a short description of the risk management process according to ISO 31000 (2009) (Figure 2).

Before any individual risk is investigated, the organization needs to establish the context in which risk management takes place. The internal and external context is considered along the context of the risk management process. Examples of the internal context are the objectives of the organization or its governance, policies, and culture. The external context contains, among others, the organization's social, legal, and economic environment. The context of the risk management process includes the process's goals, methodology, and scope. Finally, the risk criteria used during risk assessment need to be defined.

Individual risks can be assessed upon this common understanding of the context. The risk assessment is split into risk identification, risk analysis, and risk evaluation.

During risk identification, a comprehensive list of risks is created. ISO 31000 (2009) stresses that a broad scope should be used in this process step, including opportunity risks and risks outside of the organization's control. Risk identification also includes the identification of consequences tied to the risk.

In risk analysis, a more detailed understanding of the risk is developed. This entails the collection of information about the risk, the measurement of its severity, and an analysis of potential actions.

Based on the risk analysis, the risk can be evaluated by comparing it to the context of the organization, e.g., its risk appetite. This might result in the decision that the risk is acceptable, that further analysis is necessary, or that measures should be taken to modify the risk.

During risk treatment, measures are implemented to adjust the risk to the goals of the organization. For example, this could happen by removing risk sources, changing the likelihood of the risk, or sharing the risk with other parties.

The risk management process is cyclical and requires active monitoring and reviewing. Each risk should be reevaluated periodically in order to update the risk assessment and adjust the risk treatment.

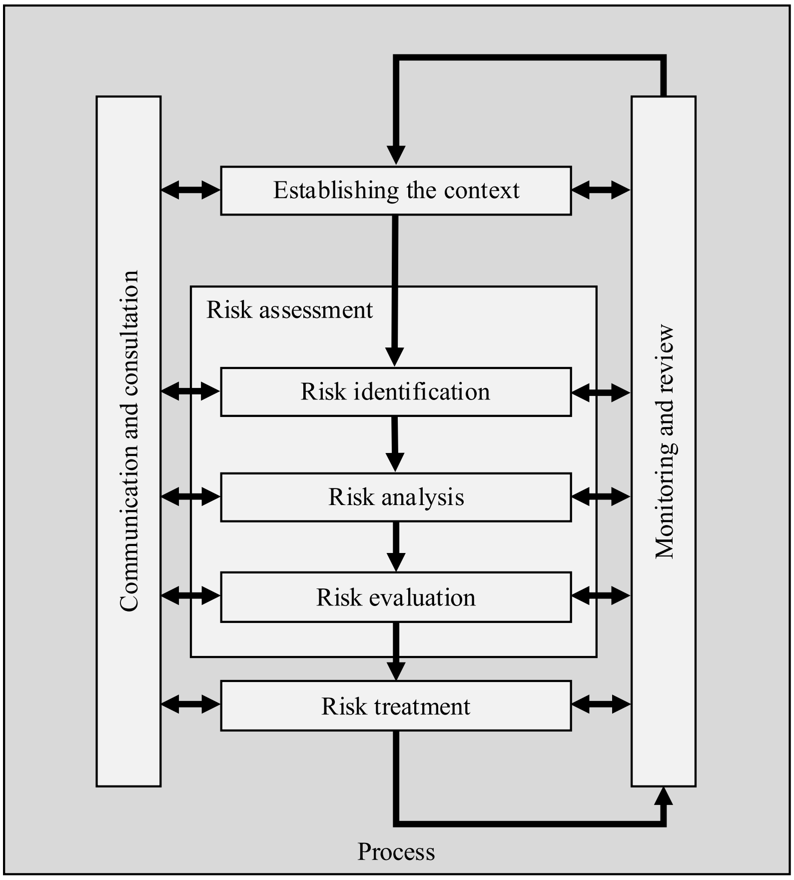


Figure 2 The ISO 31000 risk management process.[6]

2.2.2 The ERM tools

For all steps of the risk management process, formalized techniques have been developed and are widely adopted. Reviewing these tools of practitioners allows an insight into the underlying concepts that govern cognition regarding risks in organizations. All of these techniques translate knowledge from individuals within or outside the organizations into manageable items, representations of actual risks that are accessible for documentation and discussion: They provide an embodiment of the actual risks.

Two types of techniques can be distinguished: Those that focus on representing a whole set of risks and those that provide detailed representations of individual risks.

[6] The figure is based on a figure by ISO (2009).

Techniques that focus on a set of risks usually reduce the description of individual risks to a few key attributes. They allow for identifying interactions between risks and prioritizing among the risks. Among the most common approaches from this set of techniques are risk registers and risk maps.

Risk registers (also: risk tables or risk databases) are lists of all identified risks, along with selected relevant attributes. The risk register is often in the form of a spreadsheet or a database maintained by dedicated software (Hopkin and Thompson, 2022; Pritchard, 2015).

Risk maps (also: risk matrices) depict a collection of risks in a diagram, providing an overview of a whole risk portfolio (cf. Figure 3). The diagram usually places the risks along two dimensions representing measures of risk severity, for example, the likelihood of the risk event and the associated impact (Bao *et al.*, 2017; Goerlandt and Reniers, 2016).

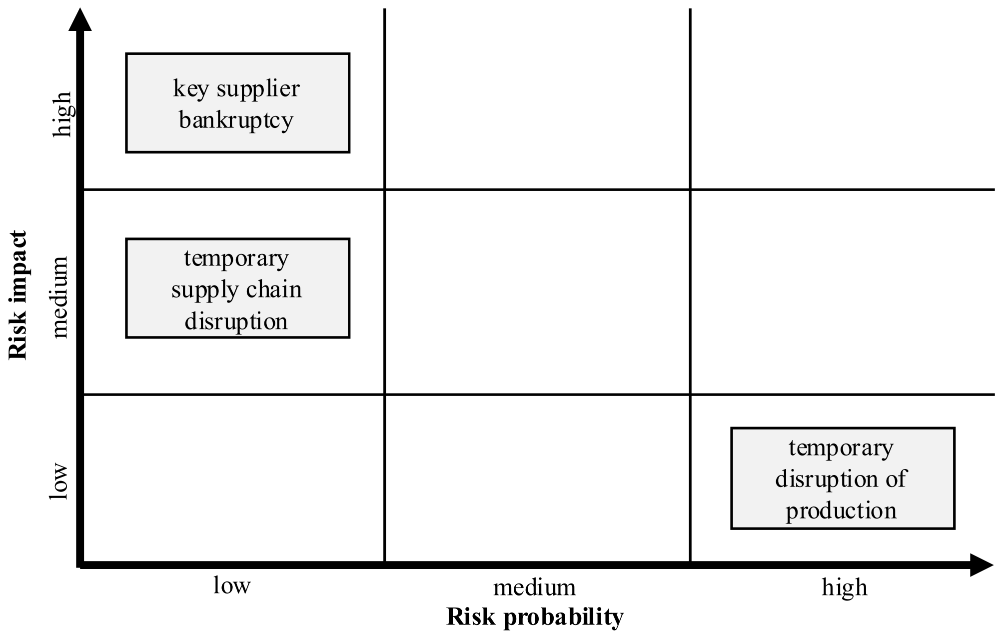


Figure 3 A simplified example of a risk map for a production plant.

Tools that provide detailed representations of individual risks are used to assist risk analysis by providing a structure to gather and organize relevant information. Information is relevant to the risk assessment if it helps to understand the severity of a risk or allows to identify actions to reduce the risk. The tools used to model individual risks provide not only a representation of the individual information and their relationship to the risk (e.g., as a "trigger" or a "consequence" of the risk under analysis) but also of the relationship between information, e.g., causal or probabilistic relationships.

A widespread tool to gather and organize information on individual risks is the **bow-tie diagram** (Ferdous *et al.*, 2012; de Ruijter and Guldenmund, 2016). A bow-tie diagram provides a graphical representation of two main aspects of risk management: preventing the risk event from happening and mitigating potential consequences of the risk events (cf. Figure 4). First, potential triggers of a risk event are collected. For each potential trigger, preventive actions are identified. For example, a development project of an industrial company might fail (risk event) because a crucial supplier goes bankrupt (trigger), leading to a breach of contract towards the customer of the company (consequences). Once triggers and consequences are established, preventive barriers (e.g., using additional suppliers) and mitigative barriers (e.g., modifying the contract with the customer) can be identified.

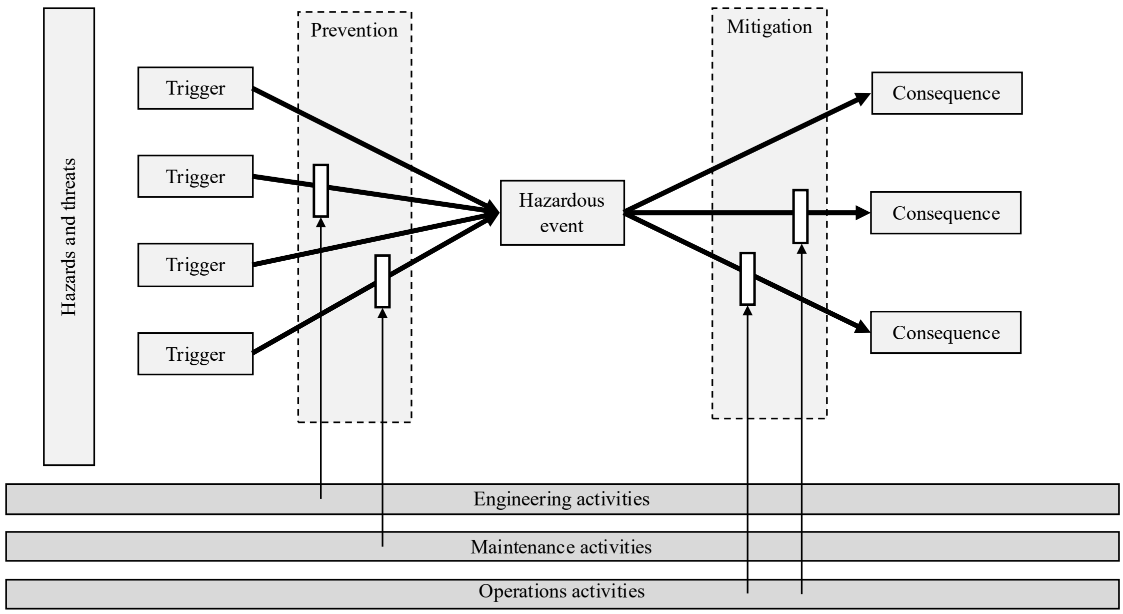


Figure 4 Example of a bow-tie diagram.[7]

A less structured visual representation of information related to risks is provided by **causal maps** (Ackermann *et al.*, 2014; Nadkarni and Narayanan, 2007). Causal maps organize information by their causal relationships to each other (cf. Figure 5). While bow-tie diagrams are fundamentally based on causal relationships as well (the risk event is

[7] The figure is based on the bow-tie diagram in Rausand (2011).

caused by triggers and causes consequences), causal maps allow for a more detailed exploration of causal relationships. For example, one potential trigger might be caused by another one, or the trigger might directly cause consequences, even if the risk event at the focus of the risk analysis is prevented from happening. Causal maps have been used to assist risk managers in developing and documenting a comprehensive understanding of risks and to identify preventive or mitigative actions (Ackermann *et al.*, 2014).

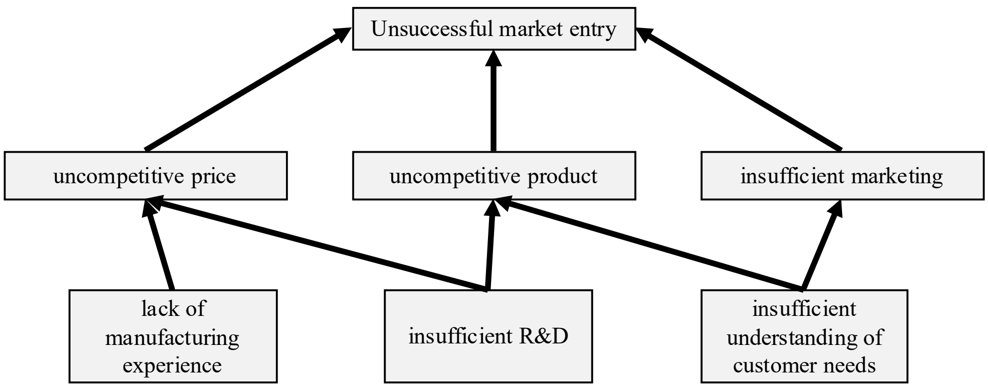


Figure 5 Example of a causal map

The causal relationships documented in a causal map can be interpreted as simple logical statements (Jensen and Nielsen, 2007; Pearl and Mackenzie, 2018): "if A, then B." In reality, relationships between events are usually less deterministic. A more reasonable interpretation would be, "if A, then B is likely." **Bayesian networks** are a tool to enrich causal maps (usually called causal models in this context) by providing probabilistic relationships instead of logical relationships: "If A is true, the probability of B being true is x percent." The probabilities that describe the relationships can either be derived from expert judgment, similar to the construction of the causal map itself, or can be computed from actual data, if available (Fenton and Neil, 2019; Jensen and Nielsen, 2007).

Once a Bayesian network is constructed, it is accessible for quantitative analysis. Given some information about the risk described by the Bayesian network, the probability of other parts of the causal map (e.g., the risk event) can be calculated. The Bayesian network becomes a quantifiable causal map (Nadkarni and Shenoy, 2001, 2004). Bayesian networks are frequently used in risk analysis and risk-related decision-making (Fenton and Neil, 2019; Weber *et al.*, 2012).[8]

The tools mentioned above for modeling risks provide structures for thinking about and discussing risks. They highlight the relationship between risks and the relationship between information about risks. Uncovering these relationships and information is usually a task that is not done by a single individual but by a group of people with diverse information and expertise to utilize all available knowledge (Surowiecki, 2004). The group can use these tools to provide a structure for a common understanding of risks. However, gathering information and knowledge distributed in a group is no trivial task (Surowiecki, 2004).

There are several well-established examples of methods that aim to facilitate the task of gathering distributed information and knowledge while avoiding introducing biases from effects within the group (Pritchard, 2015). Some of these methods aim to eliminate any direct, unfiltered interaction of the group members, as it would happen in a face-to-face discussion. Instead, the expertise of the group members is queried in a structured format.

For example, in structured expert judgment (Cooke and Goossens, 2008; French *et al.*, 2021), experts provide their subjective probability distributions regarding questions from their field of expertise. Subjective probability distributions from several experts can be combined to identify potential consensus in the expert group while providing a measure of uncertainty regarding the consensus. This method, therefore, provides quantifications for the probability of uncertain events without any direct interaction of the experts. The experts provide their expertise anonymously by default. However, the problem owner who requested the expert judgment elicitation usually knows the identity of the experts. The problem owner can, therefore, give different weights to their judgments, e.g., based on the problem owners' assessment of the experts' qualifications or the relevance of the domains or groups they represent (Cooke and Goossens, 2008).

While structured expert judgment minimizes the interaction of experts, there are also approaches to provide structured, controlled interaction between experts without resorting

[8] The second approach to model risks used in this study, constraint satisfaction networks, are not explicitly used in ERM practice. While they are able to model how individuals might think about a risk, there are, by comparison, more complex to construct and use in practice than causal maps or Bayesian networks.

to direct discussions. One such technique is the Delphi method (Linstone and Turoff, 1975), which assesses problems that cannot be precisely analyzed because the assessment needs to build upon diverse subject judgments. The Delphi method uses multiple consecutive questionnaires provided to the experts in multiple stages. Each questionnaire builds upon the answers provided by the experts in the previous stage. For example, the first questionnaire can be used to build a consensus on the scope of the problem itself and the definition of critical terms. The second questionnaire will then provide feedback on the answers of all experts to all participants. It will either build upon a consensus in the group or ask for further clarification (Linstone and Turoff, 1975). Thus, the Delphi process allows experts to exchange knowledge and revise their assessments based on the input of other experts while preserving anonymity and avoiding potential biases from direct interaction with the group. The Delphi method requires a facilitator that provides the questionnaires and that aggregates and moderates the experts' answers (Linstone and Turoff, 1975). Compared to structured expert judgment, the facilitator plays a more active part in the Delphi method, as the facilitators' decisions impact all interactions of the experts.

Of course, the use of structured communication of experts is not limited to risk management, as they apply to a wide array of fields and problems (Goossens *et al.*, 2008; Gupta and Clarke, 1996). However, the methods mentioned above are well tailored for problems arising in organizations' risk analysis, like combining expertise from diverse backgrounds and quantifying probabilities that rely on subjective judgment (Otway and von Winterfeldt, 1992). They are, therefore, standard tools in ERM (Pritchard, 2015). The Delphi method is especially beneficial for risk identification. Structured expert judgment is mainly used for risk quantification (Pritchard, 2015).

Using tools for structured communication in risk analysis highlights the importance of aggregating the knowledge of experts from diverse backgrounds without introducing biases in the process. However, it is not always possible to rely on a fully structured process, e.g., because of time and budget constraints or because more direct interaction with the experts is needed to build a shared understanding of a risk. In these cases, risk workshops are an option to combine direct interaction with a structured approach toward gathering and aggregating knowledge.

2.2.3 Risk workshops

Risk workshops are face-to-face meetings of individuals from diverse backgrounds (e.g., from different departments, levels, or functions) that aim to use the knowledge of the individual participants for a risk analysis task (COSO, 2017). Especially if a large group of participants ("more than seven or eight," Quail, 2011) is needed for the risk analysis, a structured risk workshop helps deliver results effectively (Quail, 2011).

According to Quail (2011), risk workshops are not only valuable for effectively aggregating the knowledge of a large group of experts, but they also provide a learning opportunity (as the participants themselves gain a better understanding of the risk and perspectives of experts from other backgrounds), team building (as participants get to know each other in a discussion with equal rights and opportunities to share ideas and knowledge) and education in risk management (as a structured approach to risk management is demonstrated for a specific case). Further, risk analysis often requires input from many participants, and risk workshops bring these participants together for a targeted discussion of specific problems. Finally, repeated risk workshops allow for continuous improvement in the usage of risk management tools (Quail, 2011).

There is not one definite protocol for how a risk workshop should be organized. While risk workshops are often mentioned as appropriate tools in risk analysis (cf. COSO, 2017; Fraser and Simkins, 2016; Hunziker, 2019; ISO, 2009; Tommaso, 2017), most descriptions only provide a vague outline. In the following, the procedures during a risk workshop are provided as described by Quail (2011), one of the most comprehensive guides for risk workshop facilitators.

Usually, a risk workshop will assess several risks, one after another. A risk workshop will take several hours to multiple days, with at least 40 minutes spent on any individual risk (Quail, 2011). The potential risks discussed in the risk workshop need to be defined in advance. This choice can be made before the workshop either by the sponsor of the risk workshop individually or by all participants collectively, e.g., by conducting polls or interviews. The second option reduces the probability that important or emerging risks are excluded from the workshop. Also, including the participants in the selection of risks promotes a sense of ownership toward the workshop within the group. Also, a combination of both approaches is possible, where the participants choose the risks to discuss from a pre-compiled list provided by the sponsor (Quail, 2011).

The choice of participants for the risk workshop is dependent on both the risks that are analyzed (i.e., all relevant perspectives on the risk need to be covered by the participants) and the goals of the workshop: If a decision should be made on the spot, the relevant decision maker needs to be present. If the risk workshop should serve team building, extra care is called for so that no team member feels left out. The sponsor can also include external participants if their perspective on risks would otherwise be missing (Quail, 2011).

Each discussion of a risk during the risk workshop starts with building a shared understanding of the risks within the group. To that end, the group is tasked to gather scenarios regarding the risk, like possible triggers or consequences. All participants contribute to a list of the most important scenarios regarding the risk, which is recorded and displayed during the ongoing discussion (Quail, 2011). Afterward, the magnitude of the risk is assessed by the group. This is done by an iterative process of votes and discussions, where participants record their individual assessment of the risk's magnitude and afterward explain their thoughts to the group, followed by the next vote. The process might result in a consensus within the group regarding the risk magnitude. If no consensus is found, the records of the discussions and votes at least provide an understanding of the scale of disagreement within the group and the reasons for these different perspectives (Quail, 2011).

Next, the participants are tasked with evaluating the strength of mitigants regarding the risk that is already in place or planned. Again, a vote on this question can be followed by a discussion of the reasoning of the individual participants, followed by another vote. In a similar fashion, the participants vote and discuss the probability of the risk event – considering everything learned in the discussion – and finally tolerability of risk to the organization (Quail, 2011).

The risk workshop produces specific results for every risk discussed (e.g., a consensus on the risk magnitude and probability) and a record of discussion points, information, and judgments that lead to these results. Based on these results, decision-makers can make informed decisions on how the organization should address the risks discussed in the risk workshop (Quail, 2011).

2.2.4 Biases in risk workshops

As mentioned earlier, the direct interaction of experts during the discussions of a risk workshop could harm the quality of the resulting risk assessment if potential group-induced biases are not adequately addressed. Hunziker (2019) overviews group-specific biases which are relevant to risk workshops and guides how they can be addressed. He explicitly highlights authority bias, conformity bias, groupthink, hidden profiles, and social loafing (Hunziker, 2019).

Authority bias is the tendency to give greater weight to the opinions of those perceived to have authority regarding an issue (Milgram, 1963). In the context of a risk workshop within an organization, this is especially concerning because the source of authority of a participant might not be the level of expertise but the participant's level within the organization's hierarchy. Risk workshops benefit from an atmosphere of mutual trust, where all participants feel free to offer their perspectives (Hunziker, 2019; Quail, 2011).

Conformity bias is introduced when participants adapt their assessments to comply with perceived norms within the group (Kelman, 1958). Compliance can be the conscious decision of the participants because they want positive reactions from the group ("compliance," Kelman, 1958) or they want to improve or maintain a positive relationship with the group ("identification," Kelman, 1958), or the induced behavior itself is intrinsically rewarding ("internalization," Kelman, 1958). The lack of anonymity in risk workshops is the primary source of conformity biases (Hunziker, 2019). Hunziker (2019) suggests incorporating anonymous surveys into the process or introducing outsiders into the group to address conformity bias.

Groupthink is a group's tendency to aspire to a consensus at the cost of suppressing dissent and alternative lines of reasoning (Janis, 1972). This can lead to self-reinforcing effects like polarization in the risk workshop when participants move towards more extreme positions of being either extraordinarily risk-averse or having an extreme risk appetite, as the group reinforces the more extreme positions within the group, as information that does not fit the prevalent opinions in the group is not shared (Hunziker, 2019; Shefrin, 2016). Groupthink can be mitigated by assigning explicit roles to the group participants (e.g., acting as a devil's advocate) or by changing the structure of the discussion (e.g., encouraging participants with diverging assessments to talk first) (Shefrin, 2016).

Hidden profiles refer to information that is only available to some members of the group and that is not shared with the group during the discussion (Stasser and Titus, 1985). Such information cannot, therefore, be used by the other participants for the risk assessment. Furthermore, the hidden profile is biased toward information contrary to the participants' preferences (i.e., the participants do not share information that would negatively impact their position), which excludes crucial information from the discussion (Stasser and Titus, 1985).

Finally, social loafing describes the effect that people exert less effort if they work in a group rather than individually (Simms and Nichols, 2014). A common way to mitigate social loafing is to highlight the importance of each individual's contribution (Simms and Nichols, 2014).

In conclusion, risk workshops are a standard tool in ERM to analyze risks. They are well suited to aggregate diverse perspectives on risks into a risk assessment. However, they expose the risk analysis process to group-related biases that need to be adequately understood and accounted for by the risk workshop facilitator.

2.3 Risk and organizations

Organizations are not neutral in their interaction with risks. Similar to individuals, who have a unique perception of and relationship to the risks they face, each organization has its own way of understanding and working with risks. This relationship with risks can be described in terms of culture and maturity. In the following, I provide a short introduction to the concepts of risk culture (2.3.1), ERM maturity (2.3.2), and calculative culture (2.3.3).

2.3.1 Risk culture

The previous subchapter outlined how organizations manage risks. However, the practice of risk management in an organization is insufficiently described by the techniques used for risk management. An organization also has a risk culture, which is a shared attitude towards risks, the entirety of "individual and group values and of attitudes and patterns of behavior" (Hopkin and Thompson, 2022, p. 290). COSO (2017) highlights the importance of establishing a risk culture that is universally accepted within the organization to steer the organization towards decisions that achieve the stated objectives while minimizing risks. Risk culture goes beyond a set of tools used for risk management. It includes standards and rules for behavior, compensation schemes, and norms of interaction within the organization. The totality of these characteristics places the organization on a spectrum from "risk averse" to "risk aggressive" (COSO, 2017, p. 33). Every organization needs to define its right spot within this cultural spectrum. For example, a nuclear power plant is most likely to embrace a risk-averse culture, while a hedge fund can be risk-aggressive (COSO, 2017).

No matter where an organization positions itself within the cultural spectrum, it needs to implement risk awareness (COSO, 2017; Hopkin and Thompson, 2022; Lam, 2013; Moeller, 2011). A risk-aware culture aligns the behavior of employees and the management with the organization's risk appetite. The leadership team has to make a conscious effort to achieve a risk-aware culture: Besides providing clear leadership regarding the organization's strategy and operations, it must ensure that all stakeholders are involved throughout the risk management process and provide training. Furthermore, it must avoid a blame culture, maintain clear accountability, and communicate openly about risk management issues (Hopkin and Thompson, 2022). Beyond this guidance, embedded in several risk management frameworks, Power (2020) argues for the importance of an organization's information infrastructure to sustainably impact risk culture. He determines that the appetite for knowledge of an organization, i.e., the degree to which it systematically produces information regarding itself and its environment, is a critical part of risk culture.

2.3.2 ERM maturity

A well-established risk-aware culture is a significant steppingstone toward a mature ERM. There is no agreed-upon measurement of ERM maturity, but it is often used as a proxy for the quality of an organization’s risk management practice (Callahan and Soileau, 2017; Farrell and Gallagher, 2015). Academic studies that include analysis of ERM maturity have used the appointment of a chief risk officer (C.R.O.), details from SEC filings, or simple one-dimensional scales to assess the ERM maturity of organizations (Mikes and Kaplan, 2013). Hopkin and Thompson (2022) propose a scale of four levels to classify an organization's risk maturity. The scale reaches from Naïve (unaware of the need for ERM, fragmented risk management activity focused on compliance), over Novice and Normalized to Natural (risk-aware culture, proactive ERM, risk as a competitive advantage, risk is considered in all business decisions).

2.3.3 Calculative culture

As mentioned earlier, some scholars caution that the maturity of formalized risk management practice (part of which is an explicitly stated risk culture) is not the only cultural dimension that needs to be considered when assessing the risk management of an organization: The attitude towards knowledge, its collection, and usage might be more challenging to measure but are essential aspects of how risks are managed (Aven, 2012; Power, 2020). This introduces a distinction between the explicitly expressed risk culture (e.g., “we encourage a transparent risk culture, providing opportunities for discussing uncertainties and challenges, fostering shared responsibility across all departments”) and the actual practice within the organization.

Power (2007) describes two prototypical practitioner identities that employ two different logics of calculation regarding operational risks: calculative idealists and calculative pragmatists. Calculative idealists aim to interpret all risks within the computational framework of Value at Risk. They assume that every risk can be calculated by assessing its probability and economic costs. All risk-related decisions can be based on these calculations. They are concerned about risks for which no reliable data is available and are skeptical of risk management methods that are not sufficiently robust (Power, 2007).

In contrast, calculative pragmatists do not take numbers in risk analysis at face value (c.f. Porter, 1995). They believe that numbers can only insufficiently capture the complex reality of risks. They are satisfied with relying on approximations and scoring systems instead of aiming for precise calculations. Risk management is understood as a craft rather than a science (Power, 2007).

Both types of practitioners engage in risk quantification. Still, while calculative idealists regard a fully quantified risk assessment as the ultimate (even if potentially unattainable) goal, calculative pragmatists emphasize the importance of processes and internal controls. Power (2007) notes that these two types correspond to the different approaches towards uncertainty prevalent in auditing (which is more pragmatic) and finance (that relies intensely on mathematical models).

Building upon these logics of calculation, Mikes (2009) proposes two calculative cultures. She emphasizes that many risk management tools are data-driven and analytical. Thus the culture within the organization needs to be considered in the choice and usage of these tools. Mikes (2009) distinguishes between quantitative enthusiasts, that adhere to calculative idealism and prefer quantification over judgment in risk assessment, and quantitative skeptics, that adhere to calculative pragmatism and focus on experience and judgment in risk assessment. The prevalent calculative culture leads to different practices of ERM. On the one hand, quantitative enthusiasm is consistent with an approach of "ERM by the numbers" that exclusively focuses on quantifiable risks and views ERM as a computational tool. On the other hand, quantitative skepticism is consistent with "holistic ERM" that includes non-quantifiable risks and views ERM as a "learning machine" (Mikes, 2009).

The theoretical background on the topics of risk, risk management, and the relationship between risk and organizations informs our approach to investigating risk workshop effectiveness with simulation experiments. In order to build a model that allows us to better understand risk management practice, we develop the model along with the concepts presented in this chapter. This study investigates the impact of workshop design and calculative cultures on risk workshops using agent-based simulations. In the following chapter, the methodology used for the simulation experiments is introduced. A special emphasis is given to the two modeling approaches for cognition that correspond to the two different calculative cultures.

3 Methodology

In this chapter, I explain the methodological background of this thesis. The thesis aims to answer questions regarding the effectiveness of risk management practices, specifically the effectiveness of risk workshops. To this end, we want to model discussions, as they typically happen during risk workshops, for a computer simulation experiment. In the previous chapter, I have outlined how risk workshops are used to inform decision-making regarding risks. At their core, risk workshops bring together experts that communicate with each other in a face-to-face situation regarding a risk in order to make a risk assessment that makes use of the expertise distributed among the participants. In order to simulate such an exchange in a computer simulation, two fundamental design decisions have to be made: How to model the interaction between the participants and how to model the cognition of the participants (i.e., the representation of their mental model of the risk).

The overall framework for the simulation experiment with the two-layered model (with a group layer and an individual layer) is explained in subchapter 3.1. The group interaction is modeled with an agent-based model, as explained in subchapters 3.2 and 3.3. While any agent-based model of humans requires some model of cognition, the models used in this thesis provide the agents with a complex cognition that allows them to make risk assessments based on the information available to them. For this reason, the models of cognition that are used in this thesis are explained in separate subchapters (3.4 and 3.5).

3.1 Experimental framework

This subchapter develops the experimental framework for the simulation experiment. In the simulation modeling cycle, the formulation of a research question is followed by the identification of relevant elements of the target system, the choice of a model structure, and finally, the model implementation (Barth *et al.*, 2011).

First, I define the scope of the simulation experiment with regard to actual risk workshops (3.1.1), i.e., what elements of the target system the model should include. Second, I describe the overall structure of the simulation (3.1.2), i.e., how the scope is translated into a model. Finally, I describe how the interaction of the workshop participants is implemented (3.1.3).

3.1.1 Scope of the simulation experiment

As discussed in the previous chapter, risk workshops differ regarding their specific task and regarding the precise method used to address the task. In order to model a risk workshop for a simulation experiment, a specific setting has to be decided upon. The criteria for these modeling choices are twofold: On the one hand, the simulated risk workshop should reflect a typical setting that is common in real-world risk workshops. On the other hand, the choices must be coherent with the limitations imposed on the model by the available methods.

This thesis uses an agent-based simulation approach for its simulation experiments. For an agent-based model (see 3.2), it needs to be defined who the agents are, in which environment they operate, and which objectives they follow (Railsback and Grimm, 2011). In the context of a risk workshop, these requirements translate into the following dimensions:

Agents. Participants of risk workshops differ both regarding their number and roles. Typical roles are facilitators, that organize the risk workshops and lead the discussion; the risk owner, who is responsible for the decisions made regarding the risk (hereafter called 'leader'); and the discussion participants, who are invited because of their expertise regarding the risks that are discussed or because their responsibilities within the organization are affected by the risks (Quail, 2011).

For this study, the three main roles considered in the experiment are the facilitator, who leads the workshop and controls how the exchange of the participants unfolds, the leader, who is responsible for the final result, and the discussion participants, who exchange their knowledge and form an opinion regarding the risks. As these roles have different tasks in the modeled risk workshop, their behavior has to be modeled separately. All simulation experiments in this study have nine participants in the risk workshop (excluding the facilitator).[9]

Objective. Risk workshops are used at several stages of the risk management process. They can be used to identify risks, assess risks and decide on actions to be taken to address the risks (COSO, 2017).

---

[9] See Appendix 3 for a discussion of the appropriate number of participants and a sensitivity analysis.

A simulation experiment requires a task that can be solved following pre-defined instructions. Risk assessment is an appropriate task in this regard, assuming that the information about the risk that has to be processed for a risk assessment can be described in quantitative terms. Therefore, this study focuses on the assessment of risks as the task of the simulated risk workshops. We assume that all participants share the objective of achieving the best possible risk assessment.

Environment. The environment of a risk workshop is usually determined by a facilitator(Quail, 2011). Facilitators of risk workshops can choose how the participants share information and collaborate, e.g., discussion groups, brainstorming, roleplay, or scenario analysis (Chapman, 1998; Yoe, 2019).

The most common form of groupwork during risk workshops is face-to-face discussion (Quail, 2011). Therefore, we choose this mode of collaboration for the simulation experiment.[10] In this open-form discussion, everyone can potentially speak, and everyone can listen to everything that is said.

The previously mentioned methods for group work differ in the degree of structure that the facilitator or moderator must provide to assist the participants. For example, while brainstorming is often moderated only lightly, a scenario analysis requires preparation and guidance for the participants to adhere to the method. For discussion groups, the level of structure provided by the moderator might vary depending on the needs of the group and the style of the moderator.

For this study, we choose to have the facilitator as the moderator of the discussion. The facilitator can decide who gets to talk and when to end the discussion. This role is important to identify possibilities for the facilitator to positively impact the effectiveness of the risk workshop.

Not only does the risk workshop need to be modeled, but also the task that is performed by the workshop participants.

[10] Other methods of groupwork, however, provide a more structured exchange of knowledge, e.g., the Delphi method (Linstone and Turoff, 1975). From a modelling perspective, these methods are well suited for simulation experiments as well and could be addressed by further research.

The task of the agents. The previously mentioned dimensions of risk workshops are relevant not only for a model of a risk workshop but also to design an actual risk workshop. In contrast, the structure of the risk, that is, the actual cognitive task that the group is working on, is usually not known in a real risk workshop.[11] For a simulated risk workshop, however, the cognitive task must be clearly defined. Thus, a decision must be made regarding an appropriate model that derives a risk assessment from knowledge about the risk.

Having established that we model a group discussion that has to make a risk assessment, the cognitive task of the individual participants can be described as deducting a risk assessment from their prior knowledge and the knowledge they gain during the discussion. Furthermore, the deduction of the risk assessment needs to follow specified rules in order to simulate the cognitive process within the simulation. Both Bayesian networks and constraint satisfaction networks are well suited for these requirements (see 3.5.3). Therefore, the participants in the simulation are tasked with making a risk assessment by calculating a Bayesian network (study in Chapter 4) or a constraint satisfaction network (study in Chapter 5). The task is operationalized into correctly determining whether a risk is high or low.

3.1.2 Structure of the simulation experiment

With the scope of the experiment, the overall structure of the simulation experiment needs to be developed. In order to evaluate the effectiveness of the simulated risk workshops, the outcome of the risk workshop needs to be measured against a benchmark, which in this case means the ideal outcome the risk workshop could have reached (i.e., a correct risk assessment). As discussed earlier, such an objective benchmark is often missing when real-world risk assessment practice is evaluated. Within a simulation experiment, how-

[11] For example, a typical task for a risk workshop is to assess a specific risk regarding its impact and likelihood. The participants of the risk workshop will communicate, based on their prior knowledge, and try to reach a full understanding of the risk under discussion. Even though the participants share knowledge and discuss the risk, each of them can still have a different way to translate their knowledge into a risk assessment, that can only be approximated with formal rules.

ever, it is possible to define an objective truth. For the purpose of the simulation experiment, we assume that an individual with full knowledge regarding a risk, that is, full access to all relevant information and their relationship to each other and to the overall risk, would reach the ideal risk assessment.[12] If we adopt this understanding of correctness, i.e., the risk assessment that is reached with ideal usage of all available information, this ideally achievable risk assessment can serve as a benchmark for the risk assessment produced by the risk workshop.

Figure 6 depicts the overall setup for a simulation of the discussion of one risk during a risk workshop, using a constraint satisfaction network as the cognitive architecture.

First, a risk[13] is generated, either as a Bayesian network or a constraint satisfaction network (1). The generated risk is composed of information and hypotheses about the risk and their relationships with each other. From this full risk network, an ideal benchmark risk assessment is calculated (2). Afterward, the participants are provided with a subset of information, hypotheses, and relationships. Everything from the full risk network is known by at least one participant, but (usually) no participant knows everything from the full risk network (3).

Now, the actual discussion is simulated. Participants can share their knowledge and update their individual risk network with the knowledge they receive (4). Finally, the facilitator ends the discussion, a risk assessment is decided on, and the risk assessment made by the risk workshop can be measured against the benchmark risk assessment (5).

---

[12] Such a risk assessment might still be 'correct' or 'incorrect' ex-post, depending on the definition of correctness for a risk assessment. For example, such a well-informed risk manager might have assessed the immediate threat of a major business disruption due to a global pandemic to be low in 2019. However, even though such a disruption in fact materialized, the risk manager's risk assessment could still be considered correct, as it was the best assessment possible based on all available information.

[13] Hereafter, we use the term 'risk' in the context of the simulation to refer to the task of the workshop participants, i.e., network that has to be calculated.

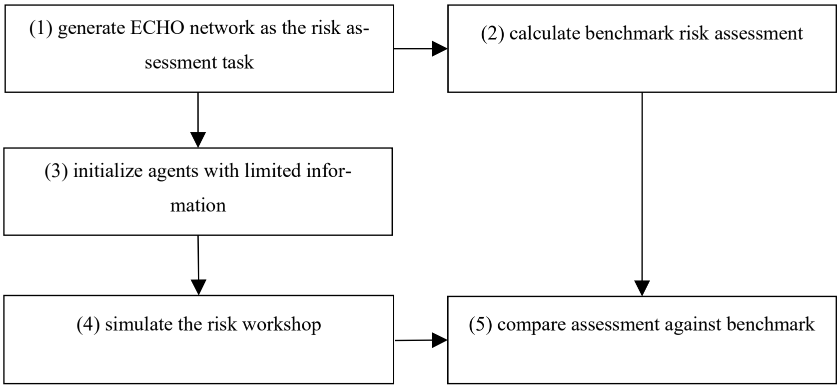


FIGURE 6 Flowchart of a simulation experiment using CSN as a cognitive architecture.[14]

3.1.3 Interaction in the risk workshops

Within the structure of the simulation experiment outlined in 3.1.2, the simulated discussion takes a central part. Here, all actors – the facilitator, the participants, and the leader – directly interact with each other in order to make use of the knowledge distributed between the participants to reach a correct risk assessment. The order of interactions is shown in Figure 7.

Each risk is discussed repeatedly and assessed by the group in several discussion rounds. Specifically, all participants initially share their assessment of the risk (4.1); then (4.2), when certain conditions are met (e.g., the group has reached a consensus), the decision maker terminates the discussion and decides on the risk assessment. If the discussion continues, (4.3) the facilitator chooses the next participant to share information (i.e., the sender). The sender shares the information (4.4). Afterward, the other participants (i.e., the receivers) update their risk assessment based on this new information (4.5).

The facilitator and the decision maker can choose different strategies for their decisions (i.e., who gets to speak and how and when to aggregate the assessments of the participants to the final risk assessment. The experimental setup allows us to measure the impact of the different strategies on the risk workshop's effectiveness.

[14] The depicted figure shows the information flow from a simulation experiment used as a preliminary study for the two studies presented in this thesis. The figure is taken from Harten (2019).

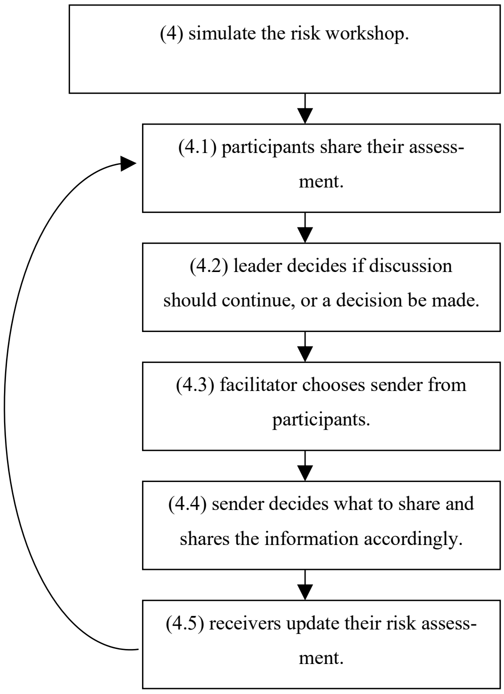


Figure 7 Flowchart of a simulated discussion during a risk workshop.

As this simulation experiment has independent actors with individual cognition and interaction between the actors, agent-based modeling is a good fit for the implementation of the simulation. In the following section, I outline how agent-based modeling has been used for similar experiments. Afterward, I discuss how individual cognition can be modeled with complex cognitive architectures.

3.2 Agent-based simulation for experiments

The simulation of systems that are characterized by interactions of independent actors is predestined for an agent-based simulation (Gilbert and Troitzsch, 2005; Railsback and Grimm, 2011). This thesis employs agent-based simulation to conduct experiments related to social interactions. Simulation experiments[15], in general, are comparatively new in this field of research. More established methods are, for example, laboratory experiments (e.g., He *et al.*, 2012; Stasser and Titus, 1985) or observational field studies. However, agent-based simulations provide significant advantages compared to other methodologies by providing a virtual testbed of reality (Hauke *et al.*, 2018; Wall, 2016). In the following, I will outline the main advantages and disadvantages of agent-based simulations compared to other methodologies and why agent-based modeling was chosen for this thesis. Afterward, I will discuss the best practices that have been proposed and employed to counter the main drawbacks of agent-based modeling.

Agent-based simulation (and simulation experiments in general) provide some valuable possibilities to researchers. First and foremost, simulation experiments provide full transparency and control to the researcher. For example, the researcher can fully control how agents behave and can also understand why agents make certain decisions. This is a solid contrast to experimental settings that observe interactions of humans, like laboratory experiments. In these cases, the researcher is limited to providing instructions to the participants of the experiments, measuring their actual behavior, and surveying ex-post the participants to gain insights into their decision-making. Using agent-based simulation, the researcher can investigate the state of each agent at any point during the simulation, e.g., to find out why a specific decision was made by the agent. Furthermore, the characteristics of the agents are known precisely. This allows for very tight control of the processes under investigation and a detailed understanding of the results (Lorscheid *et al.*, 2012).

A second strength of agent-based simulations is the possible scale of experiments. Once the model is translated into programming code, the number of experiment runs is only limited by the available computational resources. Depending on the needs of the experiment (e.g., the size of the effect under investigation), a large number of repetitions of the experiment is possible without a significant rise in costs (Lee *et al.*, 2015). This is in

[15] While simulation can have a wider meaning, this thesis uses the term to describe computer simulations, that is, using programming code to describe the underlying model and executing the program to produce simulation results. For a discussion of the term simulation, see e.g. [].

contrast to in-person experiments, where each repetition of the experiment usually comes with significant costs.

Third, agent-based simulations provide a high degree of transparency to the research community. All simulation results can be replicated with limited effort by running the simulation code. Any deviation between the reported model and the programming code can be identified ex-post by an external review of both the model description and the programming code if they are shared publicly[16].

Simulation experiments are also easily accessible for further research, building on the original experiment. If documented and published in an accessible manner, it is possible to add to the existing model, e.g., by changing the decision-making rules of the agents or providing agents with additional characteristics and immediately compare the results to the original experiment with everything else remaining unchanged.

The main challenge for agent-based simulations of social interactions, compared to other research approaches, is the need to find a satisfactory model of human action that is simple enough to model but descriptive enough to provide results that can be transferred to real social interactions (Barth *et al.*, 2011; Edmonds and Moss, 2005).

Therefore, agent-based simulations should follow best practices regarding the design of experiments (Barth *et al.*, 2011; Lorscheid *et al.*, 2012) and the transparent documentation and communication of experiments (Grimm *et al.*, 2013, 2020; Müller *et al.*, 2013) in order to avoid experiments that produce results that cannot be applied to actual social interactions or that cannot be independently verified.

[16] This transparency requires that both the model description and the simulation code is made available for external review. Within the agent-based modelling community, it is established best practice to publish both a standardized model description (Grimm *et al.*, 2010, 2020; Müller *et al.*, 2013) and the simulation code in a public repository, like CoMSES.

3.3 Agent-based simulation for group work and risk

In the previous section, we discussed how agent-based models can be used for simulation experiments. In this section, we review how agent-based models have been used on topics related to the subject of this thesis, risk workshops.

3.3.1 ABM for group work

This study aims to improve understanding of risk workshops, which are a form of group work. Agent-based models are a useful method to study group work: The agent-based model can be used to simulate interactions between the individuals in the group, represented as individual agents, as well as the environment in which the group work takes place. Agent-based models have been used to study many different aspects of group work (Fioretti, 2013; Secchi, 2015), such as the impact of individual characteristics on efficiency (e.g., Bardone and Secchi, 2017), the impact of collaboration on organizational identification (Ekmekci and Casey, 2011) or as a testbed for team processes (Lorscheid and Meyer, 2021).

Usually, the agent-based model provides the individual agents with specific characteristics or behaviors, as well as a framework for their interaction with the simulated environment. The agents' characteristics, behavior, and interactions influence the performance of the group, such as its efficiency or effectiveness, which might serve as the dependent variable in the simulation experiments (Secchi, 2015).

3.3.2 ABM for risk studies

Risk studies have used agent-based modeling for a wide array of research questions. Often, an agent-based model is used to understand a specific risk. In these studies, a system is modeled from which the risk might arise. For example, agent-based models have been used to identify effective responses to flooding events by modeling individual risk exposure (Dawson *et al.*, 2011), to model the spread of illnesses in facilities (Cuevas, 2020), or to model supply chain risks (Wu *et al.*, 2013).

However, the risk management process itself is not the focus of these agent-based models. The previous examples have shown that agent-based models are well suited to simulate group work. The two studies that are part of this thesis (Bellora-Bienengräber *et al.*, 2023; Harten *et al.*, 2022) build upon these examples by focusing on risk workshops as specific instances of group work. In order to cover the specific problems that arise when a group of individuals is tasked with assessing risks, these models need a representation of cognition that takes into account the specific characteristics of the risk assessment task. For that reason, we propose to provide the agents with complex cognitive architectures that describe cognition about risks in sufficient detail.

3.4 Bayesian networks as a cognitive architecture

To study the decision-making of individuals with agent-based networks, one needs a representation of how the individual makes decisions based on the available information. This model of the individual's cognition can take many forms. For example, a simple calculation or if-then rule might be sufficient to model an individual's cognition that chooses between well-defined alternatives. In fact, most agent-based models of social systems use highly simplified models of individual cognition (Sun and Naveh, 2004). However, while the field of simulated social sciences mainly relies on simple models of cognition, there has been ample research on how cognition can be modeled more realistically in the field of cognitive science[17] (e.g., Anderson, 1983; Thagard, 2012).

Such models of cognition are called cognitive architectures. Thagard (2012, p. 50) defines cognitive architectures as “general proposal[s] about the representations and processes that produce intelligent thought.” Cognitive architectures can be understood as analogous to computer architectures, which describe a computer with regard to its fundamental structure rather than its actual physical realization (Thagard, 2012). They provide building blocks that mimic components of cognition like memory, learning, and perception in order to build models that mimic cognitive tasks such as problem-solving, language acquisition, or vision (Thagard, 2012).

[17] Originally, the research on artificial intelligence was tightly linked to research on cognition (Langley *et al.*, 2009). However, recent advances in artificial intelligence research come from approaches that do not aim to mimic human cognition, but instead optimize models independently from any similarity to human cognition (Lake *et al.*, 2017).

3.4.1 Using Bayesian networks for ABM

This study uses Bayesian networks as the cognitive architecture of the agents in an agent-based model. It is still uncommon for agent-based models to use complex cognitive architectures for the cognition of the agents. Most agent-based models with human agents rely on a set of specific rules that the agents follow. This is in line with the general advice to keep models as simple as possible (see Edmonds and Moss (2005) for a discussion of the conflicting needs to build a model simple but also descriptive). Therefore, it is important to consider how Bayesian networks and agent-based models can be integrated into one model.

There has been some interest in the combination of Bayesian networks and agent-based models. Marcot and Penman (2019) highlight the fusion of agent-based modeling and Bayesian networks as an important approach to advance the usage of Bayesian networks. There are two main ways in which Bayesian networks can be used as components of agent-based models:

First, Bayesian networks can be used to model aspects of the environment that agents populate or the agent population itself. For example, Bayesian networks have been used for population synthesis, that is, the generation of agents that adhere to specific characteristics like demographic distributions or behavioral patterns (Borysov *et al.*, 2019; Dam *et al.*, 2011; Sun and Erath, 2015).

Second, the Bayesian network can model the agents' cognition in the agent-based model. For example, Sun and Müller (2013) use Bayesian networks to model households' decision-making regarding land-use. By modeling cognition with Bayesian networks, they can incorporate expert knowledge and quantitative data as a basis for the agent's decisions. Kocabas and Dragicevic (2013) use Bayesian networks to model how households decide on a place to live. They provide three different Bayesian networks to represent different types of households (low, medium, and high income). The Bayesian networks allow agents to make decisions about places with only limited information as evidence. Nielsen and Parsons (2007) propose a framework for formal argumentation where agents each have a Bayesian network representing their knowledge of a domain. Within the framework, agents are communicating to learn about each other's beliefs and identify a potential consensus.

The second approach mentioned is the one used for this study. By providing agents with a complex cognitive architecture, we can be more descriptive regarding the actual cognition of individuals regarding a complex decision problem. In the following, I will outline how Bayesian networks work and how they can be used to describe cognition regarding a risk assessment.

3.4.2 How Bayesian networks work

In the previous section, I have introduced how Bayesian networks have been used in agent-based models. This section provides an introduction to Bayesian networks, with a focus on how they can be used to derive probabilities of risks. To that end, I outline the fundamentals of Bayesian probability theory, how Bayesian networks are represented, and how inference can be applied to these networks.

Bayesian[18] probability theory provides a framework for making decisions based on data, with a focus on updating beliefs and decisions when new data becomes available. Before (new) data is available, a *prior probability* is assigned to a hypothesis or parameter. The prior probability represents the initial knowledge regarding the hypothesis or parameter. When new data becomes available, the probability of the data with a given hypothesis or parameter is calculated using a *likelihood function*. Using Bayes' theorem[19], the *posterior probability* of the hypothesis or parameter - an updated knowledge about the hypothesis or parameter - is calculated form the prior probability and the likelihood of the new data (Jensen and Nielsen, 2007, p. 5). For example, a doctor who assessed the probability of a patient having lung cancer might know the general probability of a patient with a cough having lung cancer. Once the doctor learns new data about the patient, e.g., if the patient is a smoker, the doctor can include this information and reach a new (posterior) probability of the patient having lung cancer (given a cough).

Bayesian networks (also called Bayes networks or Bayesian belief networks) build upon Bayesian probability theory to allow for intuitive representations of complex relationships between variables. By combining Bayesian probability theory with directed acyclic

---

[18] It is named after Thomas Bayes, who worked on conditional probabilities and after whom also the Bayes' theorem is named.

[19] Bayes' theorem is usually stated as $P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$, where $P(A)$ and $P(B)$ are the prior probabilities of A and B, $P(B|A)$the likelihood of B (given A) and $P(A|B)$ the posterior probability of A (given B).

graphs, they provide probabilistic graphical models that can be used to deduct probability distributions for hypotheses or variables that depend on many other variables (Pearl and Russell, 2000). The directed acyclic graph encodes causal relationships between variables. That is, the graph has a directed edge from node A to node B when the probability of A depends on the state of B. For example, in the previous example of a doctor assessing the cough of a patient, there would be a directed edge from [patient is a smoker] to [patient has cancer].[20] Each node of the graph represents one hypothesis or variable. Bayesian probability theory is used to calculate a probability for each possible state of each node, dependent on the state of the nodes that are considered to be causal for the state.

For example, in the simple Bayesian network provided in Figure 8, the hypothesis [Patient has cancer] is either true or false. The same is applicable for the variables [Patient presents with dyspnea] and [Patient is a smoker]. For the two variables, the Bayesian network needs a prior probability independent of the other nodes, e.g., the general share of the population that smokes or the share of patients that seek treatment with dyspnea. The probability of the hypothesis being true depends on the state of the other two nodes, e.g., it would be reasonable to assume that the hypothesis is more likely to be true when both variables are also true. The Bayesian network is fully specified with the combination of the causal graph and the probabilistic relationships between the nodes.

In the previous example, the probabilistic relationships between the hypothesis and variables are constructed from expert knowledge that relies on external information (like the share of the population that smokes). When a sufficiently large dataset is available for the variables and hypothesis under investigation (e.g., medical records for many patients that seek treatment), the probabilities can be derived from the dataset directly without any further external knowledge (cf. Kabir *et al.* (2015) for such an approach in a risk management context).

Once the Bayesian network is constructed, it can be used to infer the posterior probability of a hypothesis or variable and provide partial knowledge regarding the state of the other hypotheses and variables (Koski and Noble, 2009). For example, the Bayesian network in Figure 8 can be used to calculate the most probable diagnosis for a patient that has a cough and has recently been to Asia but has no abnormal x-ray results.

[20] Note that this uses a weak interpretation of the term 'causality' that only means to identify the most plausible direction of influence between two variables. In the example, not every lung cancer needs to be caused by the patient being a smoker.

As we have seen, Bayesian networks provide a framework to model problems that require a probability distribution for a hypothesis that builds upon partial knowledge about (many) other variables and hypotheses that are causally related to the hypothesis under investigation. This is a common problem in risk assessment: The risk analysis aims to provide a probability of some risk event using relevant knowledge related to the risk. In the following, I will discuss what a generic risk Bayesian network can look like that serves as a prototypical Bayesian network for the purpose of this study.

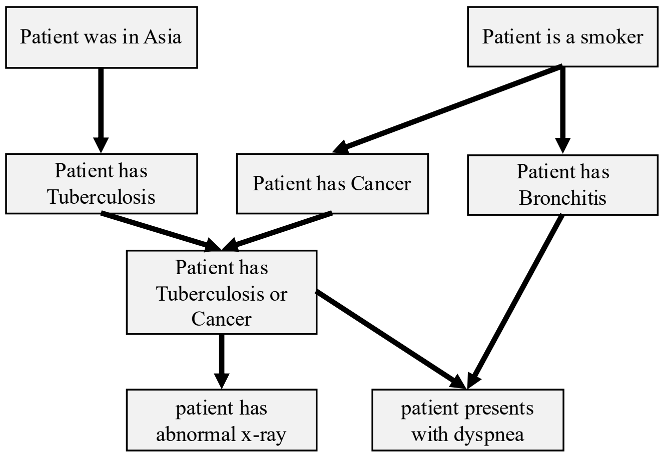


Figure 8 Example of a Bayesian network DAG: the 'Asia model.'[21]

3.4.3 A generic risk Bayesian network

In the previous section, the possibility of constructing Bayesian networks to model decision-making for problems that ask for a probability of the possible states of a hypothesis (such as true or false) was demonstrated. As discussed earlier, the task of a risk workshop can be described as such a problem: When the risk workshop is tasked to assess a risk, it must decide on the probability of the risk having certain characteristics, especially regarding its severity. In order to simulate risk workshops using Bayesian networks as the cognitive architecture of the participants, a generic Bayesian network representing a risk must be defined. In the following, I will discuss how the characteristics of the generic Bayesian network were chosen.

[21] The 'Asia model' is a popular introductory example of Bayesian networks. It was first published and discussed by Lauritzen and Spiegelhalter (1988).

In (Harten, 2019), I propose to describe the cognition of an agent regarding a risk in three layers (Figure 9).

- At the top layer, the *layer of assessment*, the risk is thought of within categories of assessment, such as “high risk” or “low risk.” These assessments are the output of the risk assessment process in the risk workshops we study.
- Below the layer of assessments, there is a *layer of hypotheses*. The assessment of the risk depends on the assessment of diverse hypotheses. For example, the assessment of the strategic risk that a new competitor might enter the market depends on the assessment of hypotheses like [our market is going to grow] or [our product is easy to copy].
- These hypotheses are assessed based on the knowledge that forms the *layer of knowledge*. For example, the existence of a patent might influence the assessment of the hypothesis [our product is easy to copy].

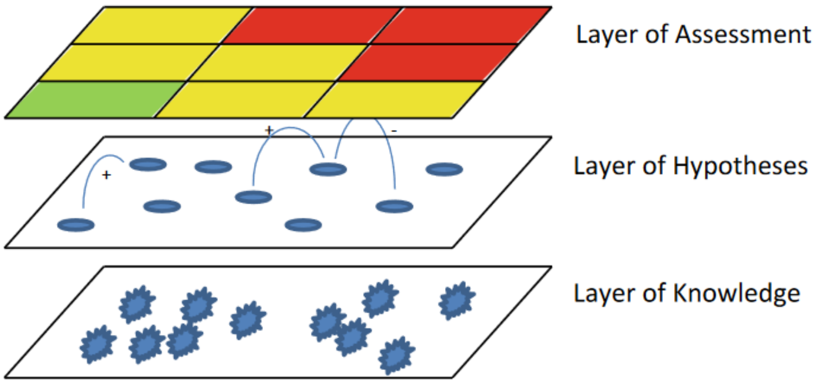


Figure 9 The three layers of the agent’s cognitive model for assessing a risk.[22]

Kabir *et al.* (2015) provide an example of a Bayesian network that is used for risk assessment in an engineering context that follows this general structure (Figure 10). It is used to assess the risk of water mains failure. In their network, the node [Aggregate Failure Risk Index] provides the overall risk assessment, it is the sole element of the *layer of assessment*. The risk assessment is derived from nodes that describe specific sources of failure, for example, the risk of failure because of lacking structural integrity or insufficient hydraulic capacity. In the proposed structure, these nodes are part of the *layer of hypotheses*, as they themselves are not representations of actual data. Rather, they themselves are derived from other nodes that represent the actual knowledge that informs the

[22] The figure is taken from Harten (2019).

risk assessment, like the water pressure, the age of the pipe, or the pH level of the soil. These nodes, finally, correspond to the *layer of knowledge*.[23]

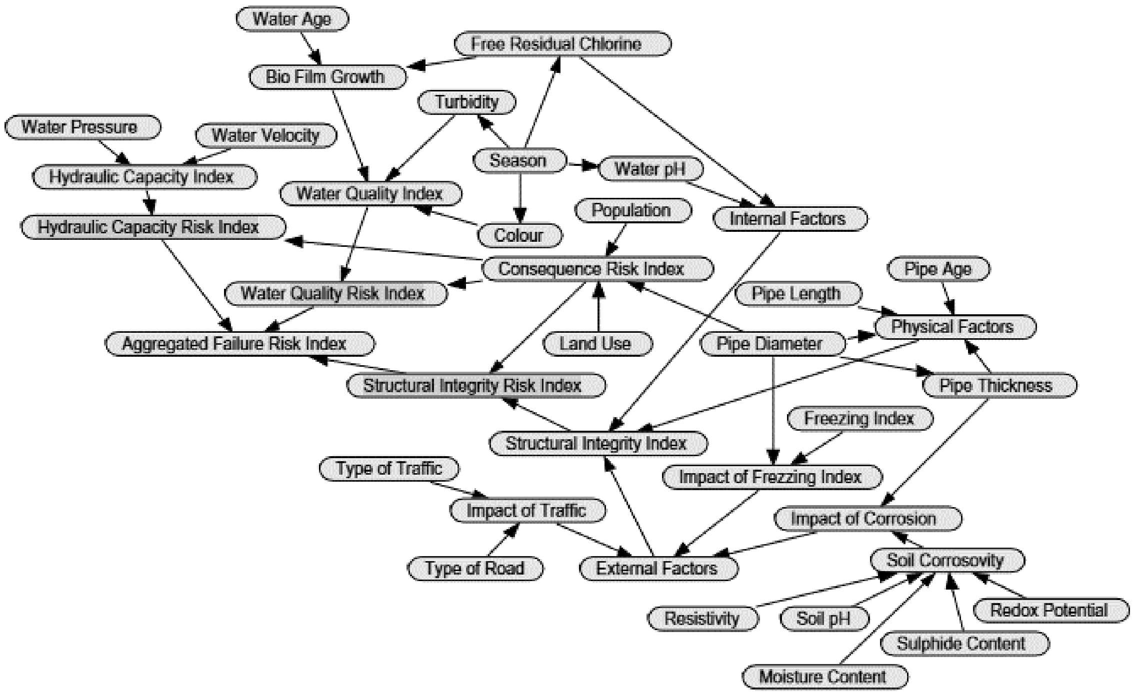


Figure 10 Bayesian network to assess the risk of a water mains failure.[24]

Generalizing this structure, we arrive at the Bayesian network shown in Figure 11. At the layer of assessment, only a single node is chosen to represent the overall risk assessment. While risk workshops often assess at least two dimensions of a risk (like the probability of the risk event and its impact), only one risk dimension is chosen for the generic risk network in the simulation study to simplify the discussion of the simulation results. As the assessment of different risk dimensions might rely on entirely distinct sets of hypotheses and knowledge, they can be described as separate problems that can be modeled using a separate Bayesian network.

Similar to the structure in (Kabir *et al.*, 2015), we generalize the layer of hypotheses into two sublayers: the overall risk assessment depends on domain nodes, which themselves depend on issue nodes. Finally, the issue nodes depend directly on knowledge, represented by 'information nodes' in the model.

[23] It should be noted that the hypotheses nodes in the Bayesian network from Kabir *et al.* (2015) are not only dependent on knowledge nodes, but also on other hypotheses.
[24] Figure taken from Kabir *et al.*(2015).

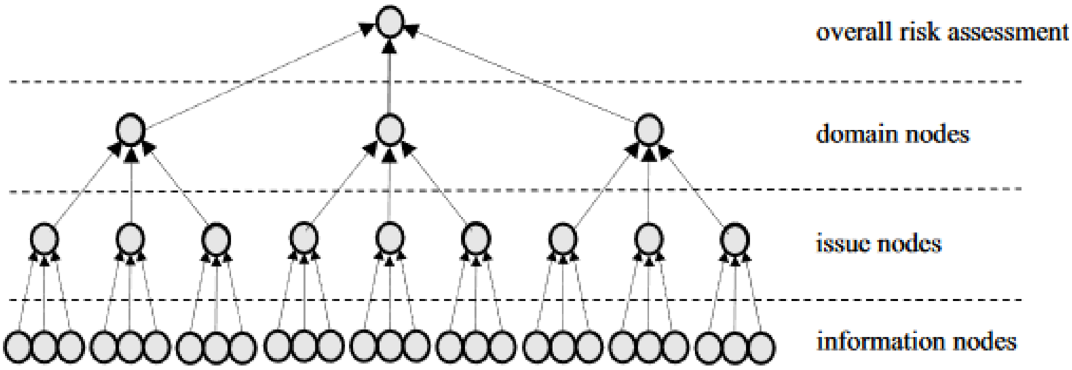


Figure 11 Generic Bayesian network of a risk.[25]

This generic Bayesian network of a risk is used in the study to represent and model the risks that are discussed by the participants in the risk workshops. The participants use the Bayesian network to reach a risk assessment based on the information provided to them: The Bayesian network serves as the cognitive architecture of the participants.

3.5 Constraint satisfaction networks as a cognitive architecture

In the previous subchapter, I have described how Bayesian networks can be used as cognitive architectures for agent-based models. Bayesian networks were chosen as cognitive architectures for this thesis because of their proximity to established approaches to model risks, as the focus of this study are risk workshops. However, there are alternative approaches to model cognition. One significant branch of research regarding cognitive architectures uses constraint satisfaction networks. They represent an alternative approach to reasoning. While Bayesian networks are built on probabilistic reasoning, constraint satisfaction networks are built on deterministic or constraint reasoning (Dechter, 2013). In the following, I will outline how constraint satisfaction networks are used in the context of cognitive architectures. There are several implementations of constraint satisfaction networks that have been used as cognitive architectures. I will focus the discussion on ECHO, a model proposed by Thagard (1989), as this model has been used for the simulation experiments in this thesis.[26]

[25] Figure taken from Harten *et al.* (2022).
[26] A discussion of several implementations of constraint satisfaction networks was published by Thagard (2000).

3.5.1 Development of constraint satisfaction networks

Constraint satisfaction networks were developed to address (partial) constraint satisfaction problems. Constraint satisfaction problems are characterized by variables for which values have to be found and constraints regarding possible combinations of values for these variables (Freuder and Wallace, 1992). For example, the variables could each represent a decision (e.g., “Which project should we prioritize?” and “Which department should get a budget increase?”). In this example, a possible constraint would be “If we prioritize project A, we must increase the budget of department X.” Possible solutions for constraint satisfaction problems are values for the variables that satisfy all constraints. Constraint satisfaction problems are often used for artificial intelligence tasks, like machine vision (Freuder and Wallace, 1992). Partial constraint satisfaction problems allow for solutions that do not fully satisfy all constraints. Instead, one tries to maximize the degree to which constraints are satisfied. There are two main reasons to settle for partial solutions to constraint satisfaction problems: Either the problem is over-constrained so that no complete solution is possible, or the complexity of the problem requires excessive resources for a complete solution (Freuder and Wallace, 1992).

The general concept of constraint satisfaction networks has been implemented multiple times as different frameworks (cf. Thagard, 2000) to fit specific use cases. A prominent framework that was developed with a focus on modeling decision-making based on arguments and observations is ECHO (Explanatory Coherence by Harmany[27] Optimization), which was developed by Thagard (1989, 1992). ECHO provides a framework for constraint satisfaction networks that describe cognitive processes as the alignment of conflicting or reinforcing units. These units represent ideas or observations of reality. It has, for example, been successfully applied to describe the reasoning about conflicting concepts or interpretations in science and judiciary (e.g., Nowak and Thagard, 1992; Thagard, 2004). While initially developed to describe the cognitive negotiation of conflicting explanations for given observations of reality, the model can and has been used to describe decision-making tasks (e.g., Frigotto and Rossi, 2015; Thagard, 2004).

[27] According to Haig (2009), the spelling “Harmany” is a deliberate tribute to Gilbert Harman, an American philosopher who coined the term “inference”.

3.5.2 Using constraint satisfaction networks for ABM

While constraint satisfaction networks are often described as a cognitive architecture (Thagard, 2012), there are only a few studies that have used constraint satisfaction networks as the cognitive architecture in agent-based models (e.g., Thagard, 2008; Wolf *et al.*, 2015).

When a constraint satisfaction network is used as the cognitive architecture in an agent-based model, that implies that the constraint satisfaction network represents the cognition of a single individual, as individuals cannot limitless access and change the beliefs of other individuals, when multiple individuals are supposed to cooperate to perform a (distributed) cognitive task within the agent-based model, that requires a model for the exchange between the individuals. This can either be done by connecting all individual networks to one shared network (as done by Hutchins (2000), using a constraint satisfaction network similar to ECHO) or by allowing the actors to communicate with each other using a separate protocol, based on their individual cognition, and updating their individual networks afterward (e.g., Thagard 2000; Frigotto and Rossi 2012).

For this study, we have chosen the latter approach (i.e., each participant has a separate constraint satisfaction network as a cognitive architecture, and communication with other participants happens in a separate layer of the model outside the constraint satisfaction network), as it is also applicable to a modeling approach using Bayesian networks. Furthermore, it allows for better control of communication between the agents, as an approach with one large, interconnected constraint satisfaction network does not provide any means to guide how individuals behave during the risk workshop. The communication framework chosen for this study is explained in Chapter 4.

3.5.3 On Bayesian networks and CSN

Bayesian networks and constraint satisfaction networks have many similarities. Both are used in the field of artificial intelligence to encode knowledge and provide mechanisms to derive decisions from that knowledge. Both have an underlying network of nodes and edges, with nodes representing variables and edges relationships. Their respective networks can serve as a graphical representation, a property that is used throughout this thesis. Finally, both can be used for reasoning under uncertainty – with incomplete knowledge about the true value of all variables, they provide algorisms that find values for the unknown variables that are consistent with what is known. For example, in the Bayesian networks and constraint satisfaction networks that represent the risk assessment task, the true value of the ‘overall risk assessment’ is always unknown but can be deducted from what is known using the mechanisms provided by the networks.

The core different between Bayesian networks and constraint satisfaction networks lies within the algorisms that represent fundamentally different kinds of reasoning. Bayesian networks are probabilistic. The key mechanism they provide to learn about the state of an unknown variable is probabilistic inference. Based on probabilistic relationships between variables (either learned from previous data or expert judgment), the probabilities of possible states of the unknown variables can be deducted. In contrast, constraint satisfaction networks solve constraint satisfaction problems. They provide values for uncertain variables that are consistent with the constraints imposed by the network on the relationship: the values they provide for uncertain variables are not to be interpreted as the most probable but as the most consistent.

As mentioned before, Bayesian networks are regularly used in fields like engineering or risk management (Fenton and Neil, 2019) as they can be used as reliable calculators for specific problems. Constraint satisfaction networks are rarely used in these fields, as they are not predictable calculators: They are designed with human cognition in mind and, therefore, best used to understand human decision-making, which is often not aligned with probabilistic inference (Nowak and Thagard, 1992; Thagard, 2004; Thagard and Verbeurgt, 1998).

3.5.4 How ECHO works

In the following, I will provide a short overview of the ECHO framework and contrast it to Bayesian networks. While the full framework is provided in (Thagard, 1989), this section will only describe features of the framework that were used for the simulation study.[28]

[28] This section is in parts taken from Harten (2019).

Similar to a Bayesian network, an ECHO network combines a graph with computational rules. The nodes of the graph, like in Bayesian networks, describe hypotheses or variables relevant to the problem under investigation. Each node has an *activation* similar to the state of a node in a Bayesian network.

While a Bayesian network requires a directed acyclic graph, the graph (i.e., the network) of an ECHO network is undirected. The edges of the graph do not represent causal relationships but instead encode the explanatory coherence of the connected nodes. There are three different possible relationships between nodes: Explanation, Contradiction, and Data.

Explanation. Explanations express positive relationships between units. If hypothesis A is logically consistent with hypothesis B, this is represented by an explanatory connection. For an ECHO network, a causal relationship between both hypotheses is not necessary—only consistency is.

Contradiction. Contradictions express negative relationships between units. If a belief in hypothesis A is not logically consistent with a belief in hypothesis B, both units are connected by a contradiction connection.

Data. A "data" relationship is established between all nodes that represent knowledge in the network and a special purpose node that is always fully activated. This relationship is supposed to make sure that stable networks must be consistent with any factual knowledge regarding the problem, as the knowledge nodes are (other than hypotheses) constantly reinforced by the factual evidence supporting them.

The ECHO network is fully described by its nodes and the connections between them. Once a network is described, all units in the network are given an initial activation. Now, the activation a of each unit is adjusted according to formula (1), until a stable state is reached (Thagard 1992):

$$a_j(t+1) = a_j(t)(1-\theta) + \begin{cases} net_j\big(max - a_j(t)\big) & if\, net_j > 0 \\ net_j\big(a_j(t) - min\big) & otherwise \end{cases} \qquad (1)$$

The formula uses the decay parameter θ to reduce the current activation of the unit (see Table 1 for the default parameter of the ECHO network). max and min are the maximum and minimum activation of a unit (−1 and +1), netj is the net input from connected units to unit j:

$$net_j = \sum_i w_{ij} a_i(t) \quad (2)$$

Here, w(i,j) is the strength of the connection between unit i and j.

As the network, over time, settles into a stable state, the activation of nodes can be interpreted as a degree of belief: A high activation of a node indicates that the node is coherent with the constraints specified in the ECHO network and the available knowledge regarding the problem under investigation. This interpretation of the network allows us to use it similarly to a Bayesian network as a cognitive architecture for a risk assessment problem. In the following section, I will describe how we construct an ECHO network for a generic risk that is comparable to the generic Bayesian network described in the previous subchapter.

Table 1 Default parameters of the ECHO framework.

Parameter	Value	Description
min	-1	Minimum activation of a unit
max	1	Maximum activation of a unit
θ	0.1	Decay parameter
a_{Start}	0.01	Initial activation of all units
Default_expl	0.05	Default weight of an explanatory connection
Default_contr	-0.2	Default weight of a contradiction connection

3.5.5 Building a generic risk CSN in ECHO

In a previous subchapter, I describe a generic risk Bayesian network that can be used as a representation of a risk for the purpose of studying risk workshops. In order to contrast two different calculative cultures, we need a generic ECHO network that is comparable to the Bayesian network discussed earlier. In the following, I describe the generic ECHO network and compare it to the generic Bayesian network (both depicted in Figure 12).

The generic ECHO network follows the same three-layered structure described in (Harten, 2019) with a layer of assessment, a layer of hypotheses, and a layer of information. As both generic networks should represent a similar problem complexity, the number of information nodes is kept constant at 27. This allows for a direct comparison of the number of information available to the participants during the workshop. While the directed acyclic graph of a Bayesian network requires a separation of the hypotheses layer into two sublayers in order to allow for connections between hypotheses nodes, these constraints on the network structure are not present in the ECHO network. Connections between hypotheses can arbitrarily be added, allowing for relationships within the group of issue nodes.

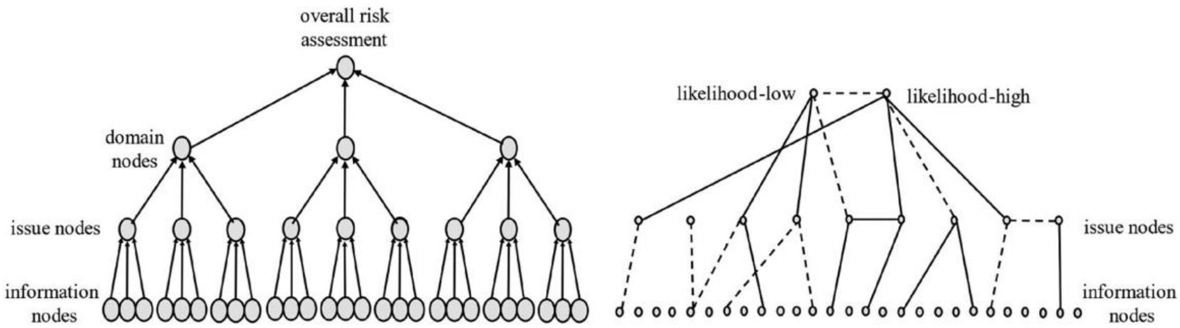


Figure 12 Comparison of the generic Bayesian network (left) and the generic ECHO network (right).

To keep the ECHO network comparable to the Bayesian network, again, only a single risk dimension is modeled. The assessment is represented by two nodes, one for the assessment 'likelihood-low' and one for the assessment 'likelihood-high.' Both nodes are modeled as contradictory to each other in order to ensure that a stable state of the ECHO network has a high activation for only one of the two possible assessments.

For a detailed description of the parametrization of the generic ECHO network, see Chapter 5 and Appendix 5. The two generic risk networks are used as the cognitive architectures of the participants, with each architecture describing a different calculative culture. While the networks provide us with a model for the cognition of the individual participants, we also need a model for the interaction between the participants during the risk workshop. The following chapter describes how the workshop as a whole is modeled for the purpose of the simulation study.

4 Effectiveness of risk workshops under quantitative enthusiasm

In this chapter, we present the results of the first study, which focuses on deriving a simulation model of a risk workshop from the idea of an ideal speech situation.[29]

4.1 Introduction

This study investigates conditions affecting the effectiveness of risk assessments in risk workshops.[30] Firms constantly adapt and transform themselves to respond to potential risks that may threaten their existence. This entails the need to correctly assess risks—a crucial task in firms' enterprise risk management (ERM) (COSO, 2017). A failure to distinguish between severe and less severe risks can generate serious detrimental consequences, up to threatening the continuation of operations. However, this assessment is not a trivial task, as decision-makers have to rely on their judgment (Mikes, 2009), and this judgment is based on information[31] that is often scattered within and beyond the organization (Neef, 2005).

[29] This chapter has been published in the Journal of Accounting & Organizational Change as: Harten, C., Meyer, M. and Bellora-Bienengräber, L. (2022), "Talking about the likelihood of risks: an agent-based simulation of discussion processes in risk workshops", Journal of Accounting & Organizational Change, Vol. 18 No. 1, pp. 153–173, doi: 10.1108/JAOC-11-2020-0197.

[30] Risk means the uncertainty about how the organization may be affected by potential events. These events may result both in positive and negative outcomes (COSO, 2017). In this chapter, for better legibility, we restrict our focus on the common focus of organizations, i.e., the focus on those risks that may result in negative outcomes (COSO, 2017). However, our modelling is applicable both for threats and opportunities. For example, when looking at interaction patterns in risk workshops, we talk about "concerned" participants; in a threats and opportunities language, a better label would be "concerned or enthusiastic" participants.

[31] We use the term "information" to refer to the participant's organized data in the context of the risk assessment task, while "knowledge" refers to the information cognitively processed and aggregated by the participants to reach an understanding of the assessed risk.

An often-used technique to facilitate the aggregation of this information is risk workshops (COSO, 2017), in which stakeholders discuss and assess both the impact and the likelihood of risks (Boholm and Corvellec, 2016). Risk assessment captures the entire process required to determine the severity of a risk after it has been identified (COSO, 2017). The severity of risk encompasses its potential impact and the likelihood of its occurrence. Risk management literature (van Asselt and Renn, 2011; Quail, 2011) suggests that the risk assessment's effectiveness, in terms of both correctly assessing the risks and the time invested to reach a decision, depends on the design and implementation of this dialogue. We investigate the design and implementation of risk workshops from that point in which the worst credible impact of a certain risk has already been established; thus, the focus is on the assessment of the likelihood of the worst credible impact of this risk. In the following, for better legibility, "high risks" and "low risks" refer to "high likelihood risks" and "low likelihood risks," respectively, and "risk assessment" refers to the "assessment of the likelihood of a risk."

Because of the difficulty of observing organizational and individual cognitive conditions in discussions (instead of merely noting its outcome) and the fact that a benchmark (i.e., the correct risk assessment and the time required to achieve it) is ex-ante absent in most risk assessments (McNamara and Bromiley, 1997), prior research has been unable to systematically disentangle different sources of (in)effective risk assessments and to describe the unfolding of the discussion over time. We address these challenges by theoretically drawing on the idea of transactive memory and Habermas' (1983) notion of the ideal speech situation. We start by suggesting that risk workshops can be conceptualized as transactive memory systems. Such a system is based on the knowledge stored in each individual's memory, the knowledge about the domain of expertise of the other individuals, and the communication about this knowledge. Transactive memory systems represent an attempt to use individuals' information by combining their expertise through a discursive process (Wegner, 1987). We then draw on Habermas' (1983) characteristics of an ideal discourse, which include free and full access to the discourse, equal opportunities to express attitudes, desires, and needs, and the absence of coercion, to define the conditions that are theoretically likely to be the most suitable to achieve the correct assessment with the least effort. Subsequently, we investigate deviations from this ideal speech situation to determine the unfolding of the risk assessment in real-world conditions.

We use agent-based modeling (ABM), which allows simulation experiments in which agents follow predefined rules to interact with other agents and with their environment (Wall and Leitner, 2021). In this study, the agents are workshop participants who communicate to assess a specific risk. ABM allows modeling the development of individual knowledge as well as its combination at the group level and related risk assessment outcomes (Secchi, 2015; Wall and Leitner, 2021). Moreover, our simulation experiments provide a correct assessment against which to evaluate the risk assessment outcome (Labro and Vanhoucke, 2007); we label this correct assessment "benchmark assessment." To define the "benchmark process," i.e., the time required to achieve the benchmark assessment, we start by simulating an ideal speech situation in which all relevant information about a risk is shared by the participants. Thereafter, we introduce more realistic scenarios representing deviations from the ideal speech situation. Specifically, we consider the effects of limits to the information transfer among participants (i.e., the receiver does only partially accept the argument of the sender because of reasons like cognitive load, time pressure, or different backgrounds), incomplete discussions (i.e., the introduction of a decision and termination approach, like voting on the risk assessment after a number of discussion rounds, instead of allowing an unlimited sharing of all information), group characteristics (i.e., unequally distributed information, hierarchical differences, and the non-recognition of the owners of expert knowledge), and specificities of the interaction patterns (e.g., prioritizing higher hierarchical positions in the discussion instead of randomly allowing an introduction of assertions).

We find that, under realistic discussion conditions, attaining the benchmark assessment is difficult, and we generate fine-grained insights on the effects of deviations from the ideal speech situation. (1) Even though the risk assessment gets stable with an increasing number of discussion rounds, limits to information transfer still can make a correct consensus unattainable. (2) In incomplete discussions, the discussion conditions are suitable either to correctly assess low or high risks.[32] An increase in the number of stable discussion rounds that are required before the leader decides about the risk assessment worsens the correct assessment of low risks' correctness. (3) Deviations from theoretically detri-

[32] We model a risk workshop that deals with a single risk. Investigating potential interdependencies in the risk assessment when discussing several heterogeneous risks in one risk workshop is beyond the scope of this study.

mental group characteristics lead potentially to higher, instead of lower levels of correctness. (4) Prioritizing participants concerned about a certain risk leads to the highest level of risk assessment correctness.

This study contributes to research and practice in three ways. First, whereas prior risk assessment research has focused on overall risks (Aven and Zio, 2014), we raise awareness of the fact that, along with an increase in duration, on average, over all risk workshops, the assessments move from an under- to an overestimation of risks. Thus, an increase in correctness in the assessment of high risks over the discussion time comes at the cost of a slight decrease in the correct assessment of low risks. Future researchers are encouraged to refine their research questions by distinguishing between the likelihood of the risks they are targeting, while firms are encouraged to allow for longer discussions if their goal is to avoid misidentifying high risks.

Second, contrary to the intuitive understanding advocated by previous risk management and group discussion literature, we show that—in the context of risk workshops—the individual characteristics of the theoretically ideal speech situation are not as ideal as presumed (Johnson and Pajares, 1996; Sheffield, 2004). For example, in terms of correctness, a decision made by the leader following the majority assessment or his/her own assessment outperforms the one made after waiting for a consensus to emerge. Firms can learn that the effectiveness of the workshop is not likely to increase after simply improving a single design component. Future research should be cautious in using the notion of ideality in discursive settings.

Finally, to the best of our knowledge, this study is the first to systematically introduce a benchmark assessment and process in risk assessment investigations. Generally, there is seldom an objectively correct assessment that serves as a benchmark (Bromiley *et al.*, 2014; McNamara and Bromiley, 1997). We overcome this limitation and avoid the commonly used singular focus on the effort required to achieve a risk assessment, instead focusing also on the decisions' correctness (Chapman, 1998; Heemstra *et al.*, 2003). Moreover, we allow to disentangle the effects of distinct deviations from the ideal speech situation; effects that are otherwise only collectively evident in the risk assessment decision (He *et al.*, 2012). While prior studies were able to account for organizational effects, like the order in which participants speak (e.g., Hiltz *et al.* (1986)), they were generally unable to capture individual information processing, like the importance an individual assigns to a received information. We model both types of effects.

4.2 Theoretical background

In the following, we provide the theoretical background on which the simulation study is built. First, we discuss how risk workshops contribute to risk assessment. Afterward, we investigate what an ideal risk workshop would look like, and which deviations are to be expected in real risk workshops.

4.2.1 Risk assessment in risk workshops

Risk workshops are instances of group discussions that form the basis for a decision made by a leader and are usually moderated by a facilitator. Relying on a group requires more effort than, for example, directly soliciting a leader's decision. However, collectively, the group is expected to make better use of the information of its individual members than the individuals alone, as the group has the chance to profit from the diversity of its members by aggregating their information on different domains (LiCalzi and Surucu, 2012; Lu *et al.*, 2012; Stasser and Birchmeier, 2003).

However, risk workshops (and, more generally, group discussions) often fail to provide reliable (risk) assessments (Hunziker, 2019; Stasser and Titus, 1985). Literature provides scattered possible explanations for these outcomes. Examples are detrimental effects because of limited information transfer due to, e.g., information overload (Paul and Nazareth, 2010) or the diversity in the participant's background (LiCalzi and Surucu, 2012). Other arguments point to incomplete discussions due to time constraints (van Knippenberg *et al.*, 2004) or group characteristics like the lack of familiarity with each other's expertise (Moreland and Myaskovsky (2000)). Moreover, the interaction patterns within the group have been found relevant (Katzenbach and Smith (2015)). For example, homogeneity and concurrence-seeking, a concept called "groupthink" (Janis, 1972), are related to suboptimal group assessment (Schulz-Hardt *et al.*, 2006). A similar effect might arise from participants being unengaged or dominating the discussion (Hunziker, 2019; Quail, 2011). While prior studies based on laboratory experiments provide clear results about individual drivers of the quality of the outcome of the discussion, they are generally unable to capture the (change of) perceptions of the individual participants and the group during discussions simultaneously affected by multiple conditions (Schulz-Hardt et al., 2006).[33] However, it is this process perspective that explains at what specific stage of the discussion process which particular decision will be made, in turn unraveling the discussion effectiveness that can be achieved under which conditions (e.g., terminating the discussion after a certain period of time, or focusing on specific participants during the discussion). This study helps to close this gap.

4.2.2 Risk assessment process: ideal conditions and deviations

We integrate a cognitive and a discursive perspective. From a cognitive perspective, we frame risk workshops as an example of distributed cognition. Distributed cognition means that groups make use of individuals' knowledge by combining their expertise (Lorscheid and Meyer, 2020). Specifically, we rely on transactive memory, a mechanism in which participants at a risk workshop learn about each other's expertise (i.e., participants build transactive memory) and then identify and combine knowledge in a discursive process. In a risk workshop, a partially differentiated transactive memory system progresses toward an integrated system. In a *differentiated* transactive memory, participants have fully disjunct areas of expertise (i.e., expertise is maximally unevenly distributed), while in an *integrated* transactive memory, all participants have the same knowledge (Wegner, 1987). Transactive memory systems have a positive impact on group performance. This impact is more likely to emerge when group members are familiar with each other's expertise and have initially distributed expertise (Lewis, 2004).

While this cognitive perspective of risk assessments focuses on the accessibility of individual knowledge to the group through discussion, a more discursive perspective complements the cognitive perspective and focuses on the design of this discussion. Habermas (1983), referring to Alexy (1978), describes the conditions of an ideal speech situation that is theoretically suited to reach a true consensus[34]. In an ideal speech situation, (1) all participants competent at speaking about the relevant topic are allowed to participate in

[33] Usually, laboratory experiment participants are surveyed before and after the discussion, while the discussion itself is simply recorded and coded. The capturing (of change) of perceptions during the discussion would disrupt the process and is generally avoided during the laboratory experiments.

[34] A consensus can be considered to be true when every competent person would agree to it (Habermas, 1971).

the discourse;[35] (2) all participants have the same chance of participating by speaking, disagreeing, and asking and answering questions, and every aspect can be discussed and criticized; and (3) all participants engage in the discussion without differences in power or other forms of coercion. The ideal speech situation is understood as a normative standard for the discussion of risks (Horlick-Jones *et al.*, 2001) that ensures that individual knowledge is properly shared and used in the group.

Real-world discussions are limited by constraints that constitute deviations from characteristics of the aforesaid ideal speech situation. Starting from Habermas (1983) and based on Handy (1986) and our summary of the literature on group discussions, we focus on four deviations:

1. **Limits to information transfer**: To reach true consensus, the speaker and listener need "[…] shared propositional knowledge, and mutual trust in subjective sincerity" (Habermas, 1982, p. 413). If these requirements are not fulfilled, a speech act might not completely convince the receiver, and, as a result, the individual's expertise on a certain risk is not fully incorporated into the assessment.
2. **Incomplete discussion**: The ideal speech situation is not limited by temporal constraints, as "no preliminary opinion [should remain] permanently withdrawn from discussion and criticism" (Habermas, 1989, p. 177). By contrast, leaders must set time limits on each risk in a workshop (Quail, 2011) and will thus enforce a termination rule, after which a leader will decide about the risk assessment.
3. **Specific group characteristics**: As there is no limit to a discussion's length in the ideal speech situation, initial differences in information access between the participants can be overcome by successively sharing information. However, if the discussion remains incomplete (i.e., it ends before arriving at a true consensus), an *unequal distribution of information* among participants may influence its outcome. Moreover, expertise might not be recognized as such (i.e., receivers have *no transactive memory*). Finally, while equal consideration of each one's arguments forms a core of the ideal speech situation, in real-world situations, *hierarchical differences* may change the acceptance of arguments.

[35] We use the term 'discourse,' which is common in Habermas' work, synonymous with 'discussion,' as used in the rest of this chapter.

4. **Specific interaction patterns**: Habermas (1989, p. 177) calls for participants to "have an equal chance to use representative speech acts" and to "have the same opportunity to use regulative speech acts, i.e., to give orders and to resist, to allow and prohibit, to make and take promises." However, in a real-world discussion, it is unlikely that participants will be equally prioritized to speak (Quail, 2011).

In line with the cognitive and discursive components of our theory, we expect that deviations from the ideal speech conditions will, ceteris paribus, reduce the risk assessments' correctness and increase the number of discussion rounds needed.

4.3 Methods

This section describes the implementation of the simulation experiment. It provides an overview of the design of the simulation and explains how the deviations from an ideal discussion are modeled for the simulation experiments.

4.3.1 Overall design

We use a simulation experiment approach, i.e., we model the reality of interest with its related processes and outcomes and combine it with an experimental design (Harrison *et al.*, 2007).[36] First, we simulate the benchmark process in line with Habermas' ideal speech situation. Second, we run four simulation experiments that model the deviations from the ideal speech situation—as described in the previous section—to disentangle the extent to which they change the effectiveness of the risk assessment. Given the importance of gaining a better understanding of the role of actors in risk management and governance (Hiebl *et al.*, 2018), we model the interaction of participants in risk workshops as the exchange of information between agents in an ABM (Lorscheid and Meyer, 2021; Wall and Leitner, 2021).

[36] The simulation code as well as the ODD+D (Overview, Design Concepts and Details + Decision) protocol are available online at www.comses.net. The protocol provides a standard for describing ABMs that include human decisions (Grimm *et al.*, 2006; Müller *et al.*, 2013). We use it for detailing the information provided in this section. The ODD+D protocol is also provided in Appendix 4.

The risk itself is modeled as a Bayesian network (Fenton and Neil, 2019; Kabir and Papadopoulos, 2019), representing both risk to be discussed and the mental model of the participants.[37] Bayesian networks are probabilistic models that describe the conditional probabilities of an event (González-Brenes *et al.*, 2016; Pearl, 2008). Combining ABM and Bayesian networks provides the two components of a transactive memory system, namely the transactive processes—reflected in the discursive interaction of the ABM's agents—and the individual memory systems—reflected in the likelihood of states represented with Bayesian networks.

4.3.2 Discussion process and risk assessment model

Each simulation experiment consists of a number of simulation runs. Each simulation run is an entire discussion of a single risk within a risk workshop and comprises five stages, as described in Figure 13. The risk structure, which is the basis for each discussion, is shown in Figure 14. The overall risk assessment (e.g., the likelihood that the introduction of a new product in the market can fail) derives from the assessment of domain-specific risks (e.g., the likelihood that competitors introduce a similar product; the likelihood that the costs are higher than the customers' willingness to pay). In turn, the domain-specific risk assessment is derived from the assessment of issue-specific risks (e.g., the likelihood that productions costs are higher than expected), and this, in turn, is rooted in the assessment of specific risk information (e.g., the likelihood that existing machines cannot be adapted to the new product). The participants' mental model is constructed analogically.

The full risk structure contains 40 nodes, comprising 27 information nodes, nine issue nodes, three domain nodes, as well as one node for the overall risk assessment. The individual participants, however, due to their diversity in backgrounds or priorities, have different risk perceptions (Sjöberg, 2000) and are initially only aware of the existence of the domains and issues related to the information they are provided with in the initialization. Before they can receive information about a certain domain or issue, they have to gain

[37] A mental model is an internal representation of a human's understanding of a system (Rouse and Morris, 1986).

knowledge of the existence of this same domain or issue through discussion with the other participants.[38]

During the discussion, information about the 27 information nodes is exchanged. The other nodes are derived from the state of the information nodes. All nodes are discrete variables in a "low," "medium," or "high" state. Each of these states is assigned a probability that represents the degree of belief that the variable is in a particular state.[39] To reflect a situation where the risk workshop needs to correctly account for a small share of critical information, we postulate in our Bayesian network that information nodes individually are ten times more likely to indicate a low than a medium likelihood and ten times more likely to indicate a medium than a high likelihood. If a participant believes that a certain information node has a "high" state (i.e., the state represented by the information node has a high likelihood), the Bayesian network will reflect this with a higher probability of the corresponding issue-, domain- and overall risk assessment nodes being in a "high"-risk state. Thus, for the same risk, participants can arrive at different risk assessments depending on the information available to them.

[38] Gaining knowledge about the existence of a new domain or issue node and getting information about the likelihood of this same node happens in different discussion rounds. When knowledge about the structure of an issue node is acquired, agents simultaneously learn about the existence of the underlying information nodes.

[39] For example, a likelihood of 100% for the "low" state signifies that the participant is absolutely certain about that assessment. A likelihood of 80% for the "low" state and, e.g., 14% to the "medium" and 6% to the "high" states indicate some uncertainty regarding the actual state of the node.

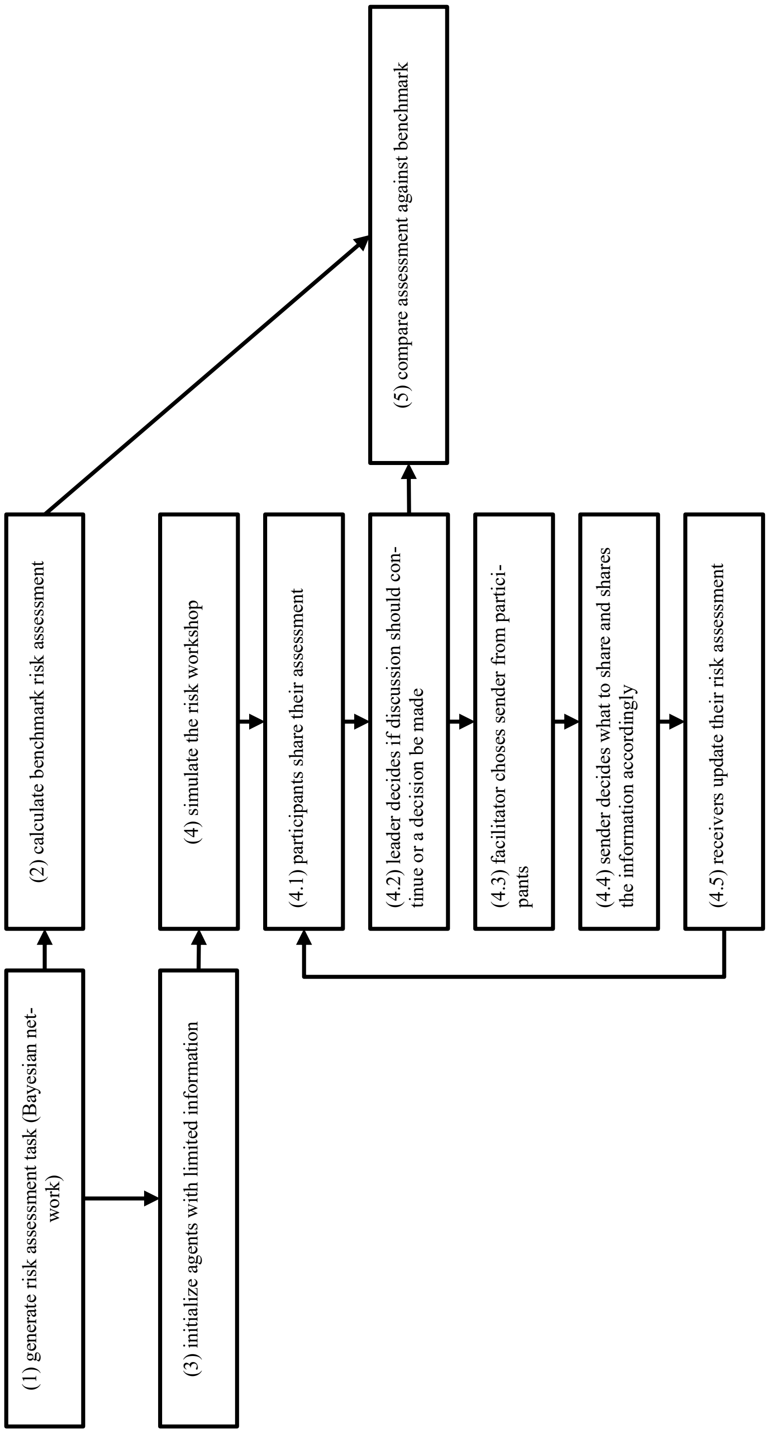


Figure 13 Flowchart of the different stages of each simulation run.

Figure 13 describes the five stages of each simulation run. (1) A risk assessment task is randomly generated according to the Bayesian network risk structure depicted in Figure 14. The task is to assess the likelihood of the risk, which is unknown to the workshop participants. (2) The benchmark assessment is calculated using the complete information from the risk assessment task. (3) Each of the participants is provided with some, but not all, information (i.e., with limited information) in such a way that each information is initially known to at least one participant. (4) Then, simulation experiments are run in line with the conditions delineated in Table 2. The risk is discussed repeatedly and assessed by the group in several discussion rounds. Specifically, (4.1) all participants share their assessment of the likelihood of the risk; then (4.2), when certain conditions are met (e.g., the group has reached a consensus), the leader terminates the discussion and decides on the risk assessment. If the discussion continues, (4.3) the next participant to share information (i.e., the sender) is chosen, and (4.4) shares the information, (4.5) followed by the other participants (i.e., the receivers) updating their risk assessment based on this new information. (5) The assessment reached by the workshop is compared to the benchmark assessment. For example, if a high risk (benchmark assessment) is assessed to be low, it is a misidentified high risk.

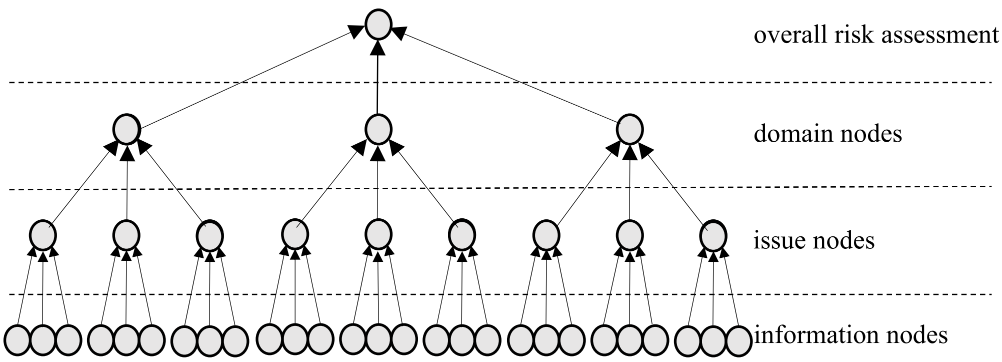


Figure 14 Graph representing the risk assessment, both as a discussion process and as an individual mental model.[40]

4.3.3 Model of the discussion

Nine participants[41] exchange information about the risk at hand. The discussion is divided into rounds. Each round consists of a sequence of actions performed by the participants

[40] The figure represents the risk structure that is the basis both for the discussion process and for each participant's individual mental model.

[41] Risk workshops can differ substantially in the number of their participants. We choose nine participants for our simulation experiment, a group size within the common range for risk workshops (e.g., Ackermann *et al.*, 2014).

(see stage 4 in Figure 13). The outcome of the discussion is influenced by how it deviates from an ideal speech situation.

4.3.3.1 Limits to information transfer

Arguments from the sender may not fully convince the receiver. In our model, when participants receive information from the sender, they will not necessarily fully discard their prior beliefs about the corresponding information node when updating their risk assessment. Instead, a receiver's new assessment of the information node is a weighted average of their prior assessment and the sender's assessment.[42] The weight that the receiver attributes to the sender's input differs across receivers and is an aggregate that, in practice, may account for factors like cognitive load, time pressure, or participant's background.

4.3.3.2 Incomplete discussions

In real-world conditions, leaders will have to determine on what basis they will make their assessment decision and when the risk workshop should end. They might choose to rely on their individual risk assessment, follow the group consensus, or rely on the majority's assessment. In terms of timing, the discussion could be stopped when a consensus emerges if the leader wants to follow a consensual assessment. Otherwise, the leader might stop the discussion when the discussion is not progressing, i.e., when the average (numerical) group assessment has been stable over a certain number of rounds (one, five, or ten).

4.3.3.3 Specific group characteristic

We focus on the impact of three group characteristics.

- **Unequal distribution of information**: Participants might not have access to the same amount of information; then, a larger share of information is provided to some participants.

[42] For example, in the ideal speech situation, the non-expert receiver will weight an expert opinion with 100%. With limited information transfer, a non-expert will weight an expert opinion with 90% and the prior belief with 10% (e.g., a prior belief of 1% in the high state of an information will turn into a $91\% = 90\% * 100\% + 10\% * 1\%$ belief after talking to an expert who assigns 100% to the likelihood of the high state).

- **Differences in hierarchy**: Information from higher-ranked participants might receive more consideration than information from other participants. Thus, the weight of the information is higher.
- **Information about each other's expertise (transactive memory)**: Participants may not know about each other's expertise (i.e., receivers lack transactive memory); thus, they cannot differentiate between expert and non-expert senders and will not weigh the information accordingly.

4.3.3.4 Specific interaction patterns

Risk workshop facilitators decide who is allowed to speak in what order, determining the interaction patterns. Using a random order as a baseline, we investigate the following interaction patterns giving priority to:

- **Concern**: The probability of being the next sender is higher if the participant's risk assessment is "high."
- **Dissent**: The probability of being the next sender is higher if the participant's assessment differs largely from the average (numerical) group risk assessment.
- **Hierarchy**: The probability of being the next sender is higher if the participant is assigned a higher hierarchical position.
- **Homogeneity**: The probability of being the next sender is higher if the participant's risk assessment is close to the average group risk assessment.

4.4 Results

Table 2 provides an overview of our simulation experiments. Figure 15 depicts the results of the simulation experiment for the ideal speech situation conditions. Specifically, it shows, per discussion round, which proportion of the simulated discussions has reached a particular type of consensus and which has not reached a consensus.[43] Before the discussion (i.e., in discussion round zero), no consensus is reached on the risk assessment in

[43] It is important to note that the Bayesian network is calibrated in a way that always results in a "low risk" or "high risk" assessment for the overall risk. This simplifies the interpretation of the simulation results. In our Bayesian network, nodes aggregate the input from three other nodes. As almost always at least some input nodes have assigned high likelihoods to

38% of the simulated discussions. The reason is that participants, at the start of the discussion, can base their risk assessment only on their limited set of information. Achieving a (correct) consensus before the discussion is driven by chance.

Moreover, we observed a tendency to initially underestimate risks (i.e., reaching a consensus but misclassifying high risks). This is due to the lack of knowledge about the existence of certain information nodes. Initially, participants are often missing information about the risk structure and therefore do not account for uncertainty regarding the probabilities of corresponding nodes (i.e., they do not yet know what they do not know). In our model, corresponding to the real-world distribution of risks, most information nodes are in the "low" likelihood state. So, participants underestimate the risk until, by learning something new about the risk structure, they become aware of their—so far unconscious—uncertainty. Therefore, in the early discussion rounds, the low risks are over-proportionally correctly identified compared to the high risks.

While, until discussion round seven, the driven-by-chance consensus drops over all simulated discussions, after this round, an increasing proportion of the discussions results in a consensus—stemming from the increased number of information shared (and, thus, a better knowledge of the risk structure and corresponding information). After at most 39 discussion rounds, all information is shared and adopted by all participants, resulting in a correct consensus for nearly all discussions.[44] The maximum of 39 discussion rounds needed is determined by the sum of the 27 information nodes, the nine issue nodes, and the three domain nodes that have to be shared to come to the overall risk assessment.

Overall, even under ideal speech conditions, it is apparent that correctly assessing a risk as a group involves many discussion rounds and is error-prone. Moreover, even if the participants reach a consensus, this consensus could be premature and wrong. Hence, the presence of a consensus is only a reliable indicator of a *correct* assessment after a large share of information has been shared.

---

the "low" or "high" states, the likelihoods assigned to the "medium" state decrease with each level of aggregation. As a result, the participants are presented de facto a binary assessment task.

[44] Due to the inherent slight imprecision of the computational framework, 100% correctness is never achieved.

Table 2 Overview of the simulation experiments

	Ideal speech situation conditions	Deviations from the ideal speech situation conditions: Limits to information transfer	Deviations from the ideal speech situation conditions: Incomplete discussions	Deviations from the ideal speech situation conditions: Group characteristics	Deviations from the ideal speech situation conditions: Interaction patterns
Experimental Conditions					
Receivers keep a part of their prior beliefs	no	**yes**	yes	yes	yes
Leader's decision approach	leader follows consensus	leader follows consensus	**leader follows own opinion, consensus, majority opinion**	leader follows majority	leader follows majority
Termination of the discussion	leader follows consensus	leader follows consensus	**after one, five, or ten stable rounds**[b]	after 10 stable rounds	after 10 stable rounds
Unequal distribution of information	no	no	no	**yes/no**	no
Receivers consider hierarchical differences	no	no	no	**yes/no**	no
Receivers have no transactive memory	no	no	no	**yes/no**	no
Interaction pattern	random	random	random	random	**priority to: concern vs. dissent vs. hierarchy vs. homogeneity**
Outcome Variables	% of correct assessments per discussion round	% of correct assessments per discussion round	% of correct assessments, avg. number of discussion rounds	% of correct assessments, avg. number of discussion rounds	% of correct assessments, avg. number of discussion rounds
Number of simulated discussions (n)[a]	1,000	3,768	3,768	7,024	3,200
Number of high / low risks	575 / 425	2,045 / 1,723	2,045 / 1,723	3,858 / 3,166	1,776 / 1,424
Results section	4.1	4.2	4.3	4.4	4.5

Note: The table presents the simulation experiments conducted, along with the experimental conditions, the outcome variables, the number of simulated discussions, the number of high and low risks (in the benchmark assessment), and the section in which the experimental findings are presented. Variables that are varied from experiment to experiment are marked in bold.

[a]A simulated discussion is the discussion of a single generated risk over several discussion rounds. In each round, a participant shares an information with the group. Each discussion was simulated for 140 discussion rounds, as this was sufficient to reach 10 stable rounds for all discussions–which is our strictest stability criterion for the termination of a discussion. Deciding on the number of simulation runs typically involves balancing computational costs and getting representative data generated by the simulation's stochastic process (Lorscheid *et al.*, 2012).

[b]A stable round is defined as a discussion round in which the risk assessment does not change from the previous discussion round. A discussion is said to have a number of stable rounds (i.e., the participants have the perception that they do not learn anymore from the discussion) if the average (numerical) group assessment does not differ more than 2% for the same number of consecutive rounds.

4.4.1 Simulation experiment 1: limits to information transfer

Figure 15 shows that when introducing limits to the information transfer, even after 78 discussion rounds—twice as many rounds as in the benchmark process—only 84% of the discussions have reached a correct consensus. As new information is not fully integrated by the receivers in their beliefs updating, senders might have to talk repeatedly about the same information to gradually increase the impact of their information on the receivers' risk assessment. At the same time, as discussion rounds continue, the classification of the group assessments becomes stable, in some cases without attaining a correct consensus. Thus, even after many discussion rounds, the unwillingness or inability to fully incorporate the sender's information impedes the achievement of the benchmark assessment.

At the top of Figure 15, the results from the first experiment representing the ideal speech situation condition are depicted (n=1,000 simulated discussions; high risks: 575 [57.5%]; low risks: 425 [42.5%]), while in the middle the results from the second simulation experiment for the limited information transfer condition are portrayed (n=3,768 simulated discussions; high risks: 2,045 [54%]; low risks: 1,723 [46%]).

The graphs in Figure 15 show, for each discussion round, which proportion of the simulated discussions has reached a particular type of consensus and which proportion has not reached a consensus to this point. At the bottom, the figure depicts the change of knowledge of the participants during the second simulation experiment. Participants are said to have knowledge about an information node if they have included the node in their mental model, and they are said to have information about the node when they have received specific information, either during the initialization or during the discussion.

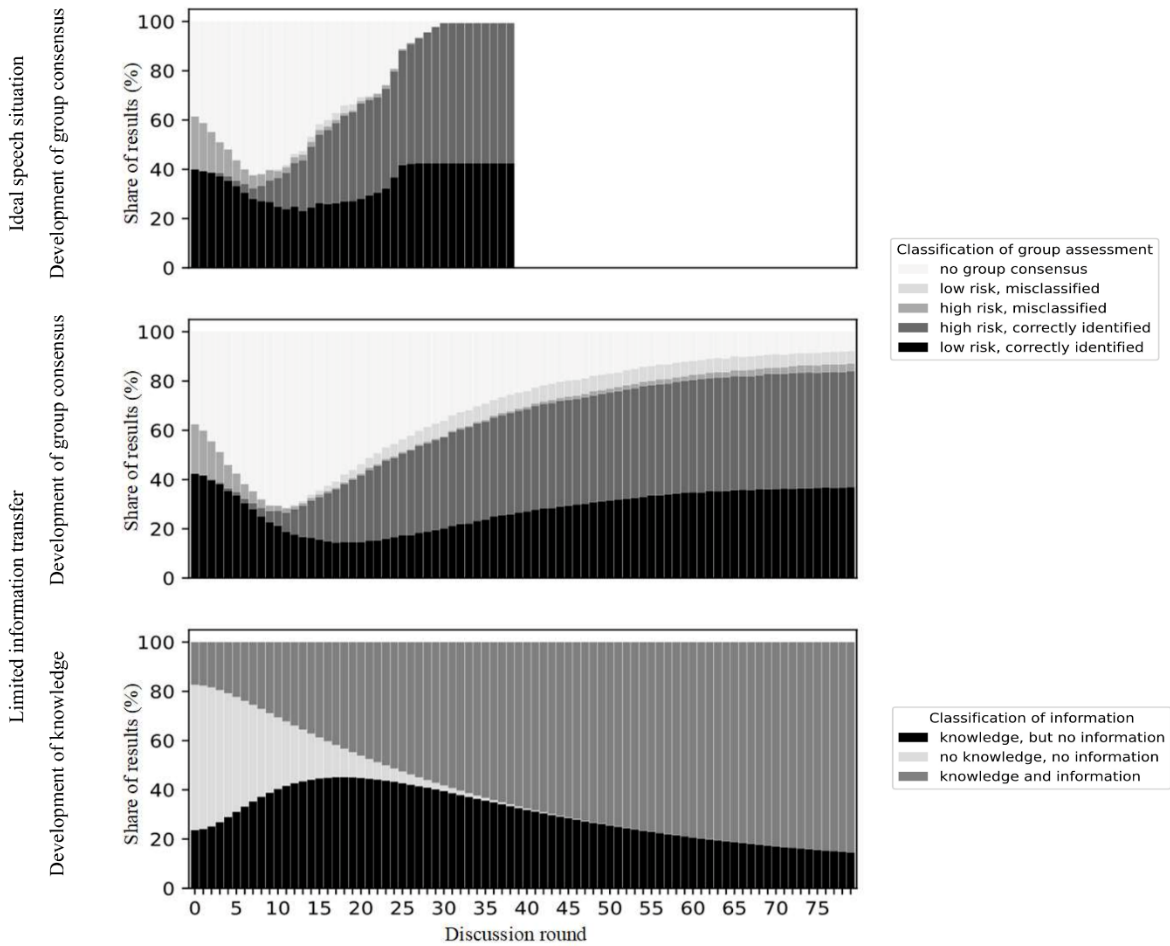


Figure 15 Development of types of group consensus after each discussion round under ideal speech situation conditions and with limited information transfer.

4.4.2 Simulation experiment 2: incomplete discussions

Table 3 aggregates the effects of three decision approaches for the leader, i.e., relying on his or her individual risk assessment, following the group's consensus, or following the majority's opinion. Leaders following their own opinion or the majority outperform a consensus requirement. For all decision approaches, we investigate what happens when the discussion is terminated after one, five, or ten stable rounds. We find that this has a clear impact on the percentage of correct assessments. Over all risks, a continuation of the discussion generally improves correctness (e.g., deciding to follow the consensus after ten stable rounds instead of five stable rounds, improves the overall percentage of correct risk assessments from 39.6% to 59.8%).

Intriguingly, correct assessments are different for high and low risks. For example, a comparison of the decision approach with the same number of required stable rounds indicates that the leader will make better decisions by following the majority if the risk is low but otherwise will improve the decision by relying on his or her individual risk assessment. Terminating the discussion as soon as a first consensus is achieved only leads to a correct assessment in 57.7% of the discussions with an actual high risk, while the same termination approach leads to a correct assessment in 97.2% of the discussions with an actual low risk.

Moreover, an increase in correctness in the assessment of high risks over discussion time comes at the cost of a slow decrease in the correct assessment of low risks. Given that firms will want to reduce the severity of the risks that they are facing, and that this severity is the product of the risk's impact and likelihood, c.p., firms will likely want to at least correctly identify the high likelihood risks and then mitigate these risks. If this holds, based on our findings, firms are encouraged to allow for longer discussions to avoid misidentifying high risks.

Table 3 Outcomes of the risk assessment under different deviations

	The proportion of correct assessments			Avg. number of discussion rounds
	All risks	High risks	Low risks	
Deviation 2: Incomplete discussion				
Stop at the first group consensus	75.8%	57.7%	97.2%	8.1
One stable round				
Leader follows their own opinion	52.5%	14.2%	98.0%	
Leader follows consensus	39.5%	0.7%	85.5%	**2.1**
Leader follows majority opinion	46.4%	1.3%	**99.9%**	
Five stable rounds				
Leader follows their own opinion	70.1%	65.3%	75.9%	
Leader follows consensus	39.6%	41.0%	38.1%	17.8
Leader follows majority	70.2%	61.2%	80.8%	
Ten stable rounds				
Leader follows their own opinion	77.9%	**82.6%**	72.3%	
Leader follows consensus	59.8%	68.3%	49.7%	33.5
Leader follows majority	**78.3%**	80.9%	75.3%	
Deviation 3: Group characteristics				
Unequal distribution of information with,				
consideration of hierarchical differences, no transactive memory	75.2%	**91.6%**	55.0%	34.5
consideration of hierarchical differences, transactive memory	73.4%	83.4%	61.8%	35.5
no consideration of hierarchical differences, no transactive memory	73.1%	80.7%	63.2%	34.6

no consideration of hierarchical differences, transactive memory	75.8%	84.4%	65.1%	36.1
Equal distribution of information with,				
consideration of hierarchical differences, no transactive memory	75.2%	82.0%	66.3%	**31.3**
consideration of hierarchical differences, transactive memory	77.1%	82.8%	70.3%	31.9
no consideration of hierarchical differences, no transactive memory	73.8%	78.2%	67.8%	31.7
no consideration of hierarchical differences, transactive memory	**78.2%**	80.9%	**75.0%**	33.5
Deviation 4: Interaction pattern				
Random choice of participants	78.3%	80.9%	75.3%	33.5
Priority given to concerned participants	**88.9%**	88.9%	**89.0%**	33.8
Priority given to participants with dissenting opinions	79.5%	**91.2%**	64.1%	33.3
Priority given to participants with higher hierarchical position	75.2%	76.0%	74.1%	32.2
Priority given to participants close to the group opinion	75.9%	70.9%	82.7%	**31.3**

Note: The table depicts the results of the third, fourth, and fifth simulation experiments and shows the percentage of risks that were correctly assessed and the average number of discussion rounds before the decision was made. For each experiment, bold values highlight the highest percentage of correct assessments per type of risk and the lowest average number of discussion rounds needed.
Deviation 2: n=3,768 simulated discussions (high risks: 2,045 [54%]; low risks: 1,723 [46%]). A stable round is defined as a discussion round in which the risk assessment does not change from the previous discussion round. A discussion is said to have a number of stable rounds (i.e., the participants have the perception that they do not learn any more from the discussion) if the average (numerical) group assessment does not differ more than 2% for the same number of consecutive rounds. If the leader follows the consensus, but no consensus is reached, the assessment is counted as incorrect.
Deviation 3: n=7,024 simulated discussions (high risks: 3,858 [55%]; low risks: 3,166 [45%]). If the information is not equally distributed, the information is distributed to the participants so that the best-informed participant knows twice as much as the second-best-informed participant, who knows twice as much as the least-informed participant. If receivers consider hierarchical differences, they weigh the sender's input according to their difference in hierarchy values: h_{low} = 0.25, h_{medium} = 0.5, h_{high} = 0.75. If receivers have no transactive memory, they do not distinguish between the input from an expert and a non-expert sender.
Deviation 4: n=3,200 simulated discussions (640 simulated discussions per interaction pattern; high risks: 1,776 [55.5%]; low risks: 1,424 [44.5%]). Compared to the ideal speech situation conditions of the first simulation experiment, in this experiment, receivers keep a part of their prior beliefs, and the leader follows the majority after ten stable rounds. When concerned participants are prioritized, the probability of being the next sender is proportional to the probability that they assign to the "high risk" state in the overall risk assessment. In a deviation from the standard sequence of actions in the simulation, in this setting, participants select the information to share with a likelihood proportional to the probability they assigned to the "high" state of the respective information node. When dissenting participants are prioritized, the probability of being the next sender is proportional to the difference between their risk assessment and the group's risk assessment. When participants are prioritized based on their hierarchical position, the probability of being the next sender is proportional to a hierarchy factor they are assigned: h_{low} = 0.25, h_{medium} = 0.5, h_{high} = 0.75. When participants close to the group opinion are prioritized, the probability of being the next sender is proportional to the inverse of the difference between their risk assessment and the group risk assessment.

This trade-off (Figure 15) is partially the result of the already discussed initial tendency to underestimate risks, as participants, at this time, lack knowledge of the full risk structure, resulting in objectively unjustified certainty (“unknown unknowns”). Hence, participants, at this time, are right with their “low” assessment, but for the wrong reason. However, as participants subsequently learn about their lack of knowledge without getting information about the likelihood of nodes, they start to overestimate the actual risk as they assign likelihoods to the new nodes. Here, participants also assign small non-zero probabilities to the “medium” and “high” states of the node for the corresponding information node. Consequently, until the participants learn the actual state of an increasing number of nodes, many participants assess the overall risk to be high and only switch to a low risk assessment as they learn about the true state of “low” information nodes.

An increase in the stability requirements is accompanied by an increase in the average number of discussion rounds required. This increase may appear trivial, but it should be noted that it is over-proportional to the number of stable rounds (2.1 for one stable round vs. 17.8 for five stable rounds vs. 33.5 for ten stable rounds). While the overall correct risk assessment only improves in a somehow linear manner, the costs of these improvements in terms of time show a steeper non-linear increase.

4.4.3 Simulation experiment 3: group characteristics

Table 3 reports the effects of a variation in group characteristics. As expected, for all risks, we observe the highest correctness (78.2%) when there are no deviations from the ideal speech for all three group characteristics; moreover, we find in the same setting the highest proportion of correctly identified low risks (75.0%). Notably, the highest share of high risks is correctly assessed when there are deviations in all three investigated group characteristics. In this condition, after the ten stable rounds required by default in this simulation experiment, the risk structure has already been learned (i.e., knowledge of the existence of the nodes has been gained); thus, the discussion focuses on the information embedded in the nodes. Here, a less optimal discussion generates noise, as the experts are not able to lower the uncertainty of other participants–as not all information is equally discussed, the hierarchically higher participants prevail over the expert, and the expertise of the experts is not recognized as such. Overall, this does not eliminate the small non-zero probabilities to the “medium” and “high” states of the nodes and leads to overestimating all risks. This is the situation in which agents are right with their “high” assessment but for the wrong reasons.

4.4.4 Simulation experiment 4: interaction patterns

Table 3 indicates that the highest correctness for all risks (88.9%) is observed when concerned participants are prioritized. Prioritizing participants that are close to the group’s opinion leads to the quickest agreement (31.3 discussion rounds) but at the cost of lower correctness. This is in line with the findings of previous literature that cautions against the concurrence-seeking inherent in the groupthink effect, specifically in risk assessments (Hunziker, 2019; Janis, 1972). Interestingly, we observe improvements when deviating from the condition of equal participation suggested by the ideal speech.

4.5 Conclusion

Risk workshops are a common technique of risk assessment and, if effectively employed, constitute a powerful instrument for risk management. However, difficulties such as defining benchmarks, disentangling different effects on the risk assessment, and capturing individual cognitive processes in discussion processes pose serious challenges to a better understanding of the design and implementation of discussion processes in risk workshops. This study responds to these challenges. It theoretically draws on the notion of transactive memory, links it to the ideal speech conditions, and investigates how deviations from this situation, likely to occur in real-world risk workshops, change the risk assessment outcomes (both in terms of risk assessment correctness and the time needed). We run five simulation experiments rooted in ABM to disentangle the effects of different deviations.

Our results provide fine-grained insights into the processes and outcomes of risk workshops. First, even though the risk assessment gets *stable* with an increasing number of discussion rounds, limits to information transfer still can make a *correct* consensus unattainable. Second, contrary to our theory and the intuition of group discussions literature, we find that increasing the number of stable discussion rounds required before conducting the risk assessment worsens the correctness for low risk. Third, we show that, for high risks, after ten stable discussion rounds, the co-occurrence of seemingly detrimental group characteristics (unequal distribution of information, hierarchical differences, and no transactive memory) leads to the highest, instead of the lowest, level of risk assessment correctness. Finally, prioritizing certain participants (namely, the ones concerned) instead of ascertaining an equal chance to speak leads to the highest level of risk assessment correctness.

Of course, this study has several limitations that could be addressed by future research. First, our analysis simulates a risk workshop that discusses one risk. While this choice was consciously made to avoid obfuscating the results with the likely effect of interdependencies across risks, we encourage future studies to use our single risk model as a baseline to investigate the effect of these interdependencies. Second, we focus on a classification task that ultimately binarily distinguishes between high and low risks. While future research might be interested in investigating the outcomes of a ternary task, we believe that our approach, at this stage, enhances the clarity of the results' communication. Third, our analysis models nine participants in the discussion. While nine is within the range of participants common in risk workshops (Ackermann *et al.*, 2014) and untabulated results qualitatively support our findings, any related choice is somehow arbitrary; future research may want to investigate the sensitivity of our findings to change in group size. Fourth, since we do not address all conceivable deviations from the ideal speech situation, future research may want to incorporate a broad range of conditions to account for factors like an increase of the limits to information transfer over time due to increasing, instead of constant, cognitive load, participants' heterogeneous motivation, or potential hidden agendas, as suggested by Bromiley et al. (2014).

Notwithstanding these limitations, we make three contributions to research and practice. First, we demonstrate that increasing the discussion rounds during a risk workshop may reduce, rather than increase, the rate of correct assessments for certain risks. Specifically, we identify a potential trade-off between the correct assessment of high and low risks. Along with an increase in duration, on average, over all risk workshops, the assessments move from an under- to an overestimation of risks. Since any improvement of the correctness of one type of risk reduces the correctness for the other, risk workshop facilitators may choose their approach for the discussion termination on this basis; for example, if the priority is on correctly assessing high risks, emphasis should be given on the longest possible continuation of the discussion (under the existing resource constraints). We contribute to research by highlighting the peculiarities in the identification of low and high risks over the duration of the group discussions. Future studies may want to include this distinction in their analyses. For example, would the results by Moreland and Myaskovsky (2000), who find a positive effect of a group member's familiarity with the others' expertise on group performance, still hold in a risk assessment setting addressing specifically high or low risks?

Second, we go beyond the idea of an ideal speech situation, as we show that this theoretical ideal might provide misleading guidance in practice:

(1) A lengthy discussion that terminates only after a high number of stable rounds does not necessarily lead to better outcomes for all types of risks. While Stasser and Stewart (1992) already concluded for their simulation of political caucuses that lengthy discussions do not necessarily lead to better decisions, we transfer their finding to a firm-based risk assessment setting, thus indicating that the specific context of discussions is not a boundary condition of this finding.

(2) A decision not based on consensual agreement does not prevent good decisions. Thus, we substantiate the conceptual claim that the final risk assessment should be based on the leader's own assessment (Quail, 2011).

(3) Rather than allowing everyone to participate in an equal way, we see that facilitators can improve the group's risk assessment more by encouraging the participation of those with concerned views. Herewith, we provide evidence for the effectiveness of an approach to countervail the concurrence-seeking of the groupthink effect in risk workshops.

Overall, risk workshop facilitators can learn from our study that an increase in the effectiveness of the workshop is likely to not be achieved by simply improving a single design component but rather requires a complete overhaul towards the theoretically ideal conditions, as shown in our benchmark process. Research can profit from our findings by using the identified conditions as a new baseline for further investigations into risk assessments. For example, we complement the work by Katzenbach and Smith (2015), who argue in favor of determining rules of interactions by providing evidence of the need to prioritize concerned participants.

Third, we contribute methodologically to the risk assessment literature by introducing a novel approach that uses ABM in combination with a simulation experiment. Herewith, we responded to the call made by Bromiley et al. (2014), who argue that studies with a known, objective risk facilitate an understanding of why and how risk assessments fail to meet expectations. While such a benchmark is usually unavailable in case studies or surveys of the risk assessment practice (McNamara and Bromiley, 1997), it can be generated by using a simulation experiment approach. Moreover, with this approach, we are able to disentangle a multitude of effects on the risk assessment in a single study, while prior studies often focus either on the aggregate effect or on a single effect (e.g., Kim and Park, 2010). Finally, it is important to note that our ABM enables us to model individual cognitive processes, including the individual's weighting of the received information and the existence of transactive memory, and the related group-level outcomes. To the best of our knowledge, this is the first risk assessment study that investigates individual cognitive processes in conjunction with organizational variables. Our modeling might serve as a steppingstone for future risk assessment investigations.

5 Effectiveness of risk workshops under quantitative skepticism

This chapter investigates drivers of the effectiveness of risk assessments in risk workshops dominated by a calculative culture of 'quantitative skepticism.' Moreover, it contrasts our findings with those of previous research that assumed the dominance of 'quantitative enthusiasm[45].'

A calculative culture captures 'attitudes towards the use and limitations of highly analytical calculative practices in an organization' (Mikes, 2009, p. 21). Mikes (2009) distinguishes between two calculative cultures, i.e., quantitative enthusiasm and quantitative skepticism. Quantitative enthusiasm considers the risk assessments made in the organization as representations of a measurable economic reality (Mikes, 2009, p. 35). Quantitative skepticism, by contrast, entails a calculative culture that regards risk assessments as learning tools supporting the holistic formation of judgments incorporating difficult-to-quantify information (Burchell *et al.*, 1980; Mikes, 2009; Power, 2007). The resulting risk numbers are seen as trend indicators, rather than a full account of reality (Mikes, 2009).[46]

We expect that an organization's predominant calculative culture[47] affects the unfolding of discussion processes for risk assessments and, thus, their outcomes. At present, the scant literature investigating the effectiveness of risk assessments in an organization's risk workshops does not explicitly address the effects of different calculative cultures and implicitly assumes a calculative culture of quantitative enthusiasm. To the best of our knowledge, this is the first study that contrasts the drivers of the effectiveness of risk

---

[45] This chapter has been published in the Journal of Risk Research as: Bellora-Bienengräber, L., Harten, C. and Meyer, M. (2023), "The effectiveness of risk assessments in risk workshops: the role of calculative cultures", Journal of Risk Research, Routledge, Vol. 26 No. 2, pp. 163–183.

[46] These calculative cultures are embedded in broader enterprise risk management (ERM) practices. While quantitative enthusiasm is a cornerstone of an 'ERM by the numbers,' quantitative skepticism is a descriptive component of a 'holistic ERM' (Mikes, 2009, p. 35).

[47] We consider these two calculative cultures as two extreme points of a continuum, similar to Mikes' (2009) suggestion that the related ERM models might be 'different stages in the evolution of risk management' (Mikes, 2009, p. 37).

assessments depending on the predominant calculative culture. Specifically, we expect that the design and implementation of risk workshops—intended to enhance their effectiveness—should be different depending on the predominant calculative culture.

A risk workshop is effective when it correctly assesses the considered risk, ideally minimizing the time needed to do so (van Asselt and Renn, 2011; Harten *et al.*, 2022; Quail, 2011). The correct assessment of risks[48] is a non-trivial task for organizations. The ability to distinguish between high and low risks is vital for any business. Nevertheless, such an assessment can be challenging as related information[49] is often distributed inside and outside the organization (Neef, 2005). Risk workshops are a common tool for aggregating information about risks (COSO, 2017). During risk workshops, participants discuss risks and derive an assessment of their impact and likelihood (Boholm and Corvellec, 2016). To improve comparability with prior research (Harten *et al.*, 2022), we focus this study on the assessment of a risk's likelihood.[50]

Previous research has struggled to systematically disentangle different drivers of effectiveness in risk workshops. A notable exception is Harten et al. (2022), who conceptualize risk workshops as transactive memory systems. Transactive memory systems combine the expertise of individuals by accessing their individual information in a discursive process (Wegner, 1987). Thus, transactive memory systems include individual knowledge, knowledge about who knows what, and communication to access each other's knowledge. Within such a theoretical framework, a calculative culture determines how individuals process information to form risk assessments.

In Harten et al. (2022), although not explicitly stated, the underlying conceptualization of information processing and formation of assessments is akin to what has been described

---

[48] 'Risk' means uncertainty about how certain events, when they occur, may affect the organization. These events can have positive and negative outcomes (COSO, 2017). In this chapter, similar to Harten et al. (2022), we restrict ourselves to those risks that may result in negative outcomes (COSO, 2017). However, our modeling applies to both threats and opportunities.

[49] In the following, 'information' refers to the participant's organized data in the context of the risk assessment task, while 'knowledge' refers to cognitively processed and aggregated information that enables participants to reach an understanding of the assessed risk.

[50] Our modeling can be also used in future research to incorporate the impact dimension of risk assessments.

above as quantitative enthusiasm. Indeed, in their model, risk workshop participants exchange information and gradually update their probabilistic beliefs about risks. This results in an overall quantitative risk assessment, which measures the overall likelihood of the risk discussed. This raises the question of the extent to which their results can be generalized when considering the boundary condition of the predominant calculative culture. In other words, would Harten et al.'s (2022) results also hold under a calculative culture of quantitative skepticism? In the latter case, risk workshops provide an opportunity to exchange risk-relevant facts and their effects concerning an organization's risks, thereby improving the understanding of the overall risk environment. We expect that the difference in the predominant calculative culture is crucial for a better understanding of the effectiveness of risk assessments in risk workshops. For example, Harten et al.'s (2022) observation that longer stagnation phases in discussions may indicate a correct assessment of risks might no longer be justified. In a calculative culture of quantitative skepticism, risk assessments might remain stable for a longer time, even when new information is presented. New information that is largely consistent with the participants' previous beliefs is absorbed instead of slightly changing the overall risk assessment.

To investigate our research question, we use agent-based modeling (ABM). We simulate how risks are assessed in risk workshops. The simulated agents are the participants who—without any hidden agenda (i.e., without own undisclosed objectives)—exchange information to reach a risk assessment for one specific risk. ABM provides a computational laboratory for controlled experiments where agents act according to predefined rules in a clearly defined environment (Wall and Leitner, 2021). Harten et al. (2022) already leverage this method to address the challenge of investigating cognitive processes and the absence of an objective benchmark for actual risk assessments (McNamara and Bromiley, 1997). ABM allows to model the development of an individual's knowledge during the discussion and the group's utilization of individual knowledge to reach a risk assessment (Secchi, 2015; Wall and Leitner, 2021). Specifically, ABM permits the representation of diverse types of information processing and judgment formation of individuals, thereby incorporating different calculative cultures.

While Harten et al. (2022) model agents' information processing and the formation of risk assessments using a Bayesian network, we use ECHO, a constraint satisfaction network (CSN) (Thagard, 2012). Harten et al.'s (2022) cognitive architecture of agents represents a calculative culture of quantitative enthusiasm, as Bayesian networks use probabilities as input and provide a probabilistic risk assessment as output, which can be interpreted as the measurement of the underlying overall risk. ECHO models, by contrast, represent a calculative culture of quantitative skepticism as they encode individual risk assessments as coherence-based relationships typifying qualitative mental models in the cognitive architecture of agents. Their output can be interpreted as a holistic judgment (vs. a probabilistic measurement of economic reality) that incorporates difficult-to-quantify information. To enable a direct comparison of the results, we follow Harten et al. (2022) in our basic experimental design.[51]

Compared to a numerically generated benchmark, we investigate scenarios that are more realistic. We simulate the effects of (1) limits to information transfer within the group, (2) incomplete discussions involving various approaches to terminate the discussion and to make a decision, (3) group characteristics like information distribution and hierarchical relationships, and (4) rules specifying interaction patterns (e.g., prioritizing participants who are concerned about a specific risk). As we conduct these experiments with agents who have a calculative culture of quantitative skepticism, we can assess which of Harten et al.'s (2022) results, rooted in a calculative culture of quantitative enthusiasm, still hold and which, in turn, are associated with the change of the calculative culture.

We find that the type of calculative culture predominant in an organization matters for the effectiveness of risk assessments in risk workshops. Indeed, some of the drivers of a risk assessment's effectiveness in risk workshops are different from those found by Harten et al. (2022). Concerning the development of the discussion over time, we document an overall improvement of risk assessments for both high and low likelihood risks. However, we regularly observe sudden and seemingly unpredictable changes in participants' and groups' risk assessments when new information overturns the previously sta-

[51] Like Harten et al. (2022), we also investigate the design and implementation of workshops from the point of view where the worst credible impact of a particular risk is clear. So, the assessment focuses on the likelihood of the risk's worst credible impact. Subsequently, to ensure clarity, 'high risks' and 'low risks' refer to 'high likelihood risks' and 'low likelihood risks,' respectively, and 'risk assessment' refers to the 'assessment of the likelihood of a risk.'

ble beliefs of the participants. The exchange of critical information can rapidly shift assessments, questioning the stagnation of a discussion as a criterion for ending it. At the same time, consensus serves as a good indicator of the risk assessment's correctness. We find that path dependencies, given quantitative skepticism, are characteristic of discussion processes, i.e., what matters for the risk assessment is not only *what* information is exchanged but also *when* it is exchanged. Initially, new information has the potential to overturn the participant's mental model, which contains little information at this early stage. Finally, the prioritization of concerned participants only results in the highest level of risk assessment correctness for high risks, while hierarchical differences among participants do not negatively affect the correct assessment of risks.

This study makes at least three contributions to theory and practice. First, to the best of our knowledge, our study is the first to show that different calculative cultures result in different drivers of risk assessments' effectiveness. The identification of the distinction between the calculative cultures of quantitative enthusiasm and quantitative skepticism, respectively, is important as a boundary condition when designing and implementing risk workshops aimed at the highest possible risk assessment correctness. The reason for this, among others, is that risk workshops in different organizations require different durations, different termination rules, and different treatments of hierarchical differences and dissenting participants.

Second, despite these differences, we also conclude that a few drivers of the effectiveness of risk assessments are resistant to changes in the dominant calculative culture. For example, the trade-off between correctly identifying low and high risks as well as the negative effect of unequal information distribution within the group is independent of the calculative culture being quantitatively enthusiastic or skeptical. This points to the ubiquitous importance of considering these process characteristics and drivers when facilitating a risk workshop.

Third, we introduce a novel methodological and conceptual approach to risk literature that incorporates the way in which (i.e., *how*) risk information is processed and used for risk assessments in ECHO networks.

5.1 Theoretical background

In the following, we outline the theoretical background of the study. First, we provide an overview of the role of risk workshops in risk assessment. Afterward, we discuss the concept of calculative cultures in the context of risk management. We link the concept of calculative cultures to specific cognitive architectures and introduce constraint satisfaction networks as an alternative to Bayesian networks.

5.1.1 Risk assessments in risk workshops

Risk workshops are used to assess the impact and likelihood of risks (Boholm and Corvellec, 2016; COSO, 2017). The workshops are discussions moderated by a facilitator, and they enable a leader to decide to use the outcome of a workshop. The group can use its participants' diverse backgrounds by aggregating their individual knowledge and thereby reach better decisions than the participants would have reached on their own (LiCalzi and Surucu, 2012; Lu *et al.*, 2012; Stasser and Birchmeier, 2003).

Therefore, risk workshops can be framed as distributed cognition (Harten *et al.*, 2022). That is, the cognitive task is not performed by individuals in isolation but by a group as a whole, using the cognition and knowledge of all participants (Hauke *et al.*, 2018). Our model implements transactive memory, as it allows participants to learn about and afterward use the knowledge of other participants.

Merging this cognitive perspective with a discursive perspective, Harten et al. (2022) identify several potential drivers of discussions' effectiveness in risk workshops from the literature. These drivers encompass (1) the effects of limits to information transfer within the group (i.e., knowledge of other participants' knowledge when integrating new information instead of discarding previously held beliefs), (2) incomplete discussions (i.e., the rules by which it is decided to end the risk workshop instead of continuing the discussion), (3) group characteristics (i.e., information distribution, consideration of hierarchical relationships, and knowledge about each other's fields of expertise), and (4) the interaction patterns applied in the group (i.e., the order in which participants are allowed to talk).

5.1.2 Calculative cultures in risk management

Recently, corporate culture has received increasing attention as a key factor influencing risk management and its effectiveness. The term risk culture is used in different ways, among others, as ‘the shared preferences towards risk and uncertainty’ (Pan *et al.*, 2017, p. 2328) or, more specifically, as a subset of organizational culture, specifying how ‘organizations think about, know, process and act upon risks and uncertainties’ (Power, 2020, p. 45). In this study, we specifically focus on the latter and address the effects of different calculative cultures (Mikes, 2009) that respectively represent different attitudes towards calculative practices.

Previous research identified two different calculative cultures in the context of two distinct approaches to enterprise risk management (ERM) in organizations.[52] Mikes (2009) studied ERM practices at two banking organizations and contrasted their ERM models as ‘ERM by the numbers’ and ‘holistic ERM.’ These different practices stem from different corporate governance pressures (Power, 2007).

First, the ERM-by-the-numbers approach to managing risks aims at measuring the impact thereof on shareholder value. This leads to an emphasis on quantifiable risks, the impact of which on shareholder value is calculable. The overall risk portfolio is described as an aggregate and can be compared to the organization’s risk appetite. In this manner, ERM contributes to the overall performance measurement. According to this perspective of risks, the calculated values for risks are the decisive output of the risk management process. The focus is on improving the quality of these calculated values by improving risk models. Obviously, this approach is ill-suited when dealing with hard-to-quantify risks. Mikes (2009) labels the calculative culture embedded in this type of ERM as ‘quantitative enthusiasm.’

Second, the holistic ERM is not directly aimed at shareholder value but at achieving the organization's strategic objectives. So, ERM focuses on the identification of what it is that puts the achievement of those objectives at risk. While quantifiable risks are still relevant from this perspective, it also takes hard-to-quantify risks into account. ERM is not used to calculate overall risk exposure but to learn about the conditions under which the organization runs. The focus is less on precision and more on inclusiveness when understanding the overall risk environment. Mikes (2009) labels the calculative culture embedded in this ERM type as ‘quantitative skepticism.’

[52] See also Arena et al. (2011) for empirical data on the interplay between ERM implementation and risk culture.

The scope of the two ERM models obviously overlaps. The same risk can potentially be addressed by using either approach to ERM. Nevertheless, as both calculative cultures derive from different objectives, the reasoning about the risk will be different. We, therefore, discuss how different approaches to cognition represent these different calculative cultures in risk assessments.

5.1.3 Cognitive architectures of different calculative cultures

When simulating processes that, like risk workshops, involve human cognition, one must choose the most appropriate model of the cognition of the people involved. This requires the choice of cognitive architecture. A cognitive architecture is a description of how information (including knowledge or beliefs) is stored in memory, how this memory is structured (i.e., the relationship between elements within the memory), and how the memory is processed (i.e., how it is utilized to learn or to reach conclusions) (Langley *et al.*, 2009; Thagard, 2012). Two architectures used to investigate how humans make causal inferences are explanatory coherence and Bayesian networks (Thagard, 2004).

5.1.4 Bayesian networks as a cognitive architecture

The ABM presented by Harten et al. (2022) to model risk workshops uses Bayesian networks as the cognitive architecture of the agents. Bayesian networks model causal relationships between nodes that represent variables of interest (Pearl and Russell, 2000). Each node is associated with a probability value, which represents the degree of belief regarding the corresponding variable's state (Thagard, 2004). The connections between the nodes represent causal links. Each node's probability value is linked to connected nodes by dependent probabilities (e.g., 'if new competitors enter the market, how likely is it that they will target the same client segment'). The probability values can either be deduced by logical reasoning or by inference from observations of reality.

Given information about the true state of some of the nodes allows deducing the probability of all remaining nodes. The possibility of calculating specific probability values from causal relationships and limited information has made Bayesian networks a popular tool in risk assessment in particular and in artificial intelligence systems in general (Fenton and Neil, 2019). Therefore, Bayesian networks are often used as a calculative tool to process information.

5.1.5 Explanatory coherence as a cognitive architecture

The theory of explanatory coherence was explicitly developed to explain why and how humans acquire certain beliefs. Among others, the theory has been implemented in the computational model ECHO (Thagard, 1989). Like Bayesian networks, ECHO networks are built from interconnected nodes. The connections describe symmetrical relationships between the nodes (e.g., 'our product has a high profit margin' is coherent with 'new competitors are attracted to our market'). Where Bayesian networks rely on specific probability values and allow precise calculations, ECHO is derived from basic principles[53] and is usually employed without adjusting concrete weights or otherwise providing numeric parameters to the network. An ECHO network can, therefore, usually be fully documented by a graph depicting the nodes and the relationships between them. Relationships between nodes are either explanatory (i.e., a 'high profit margin' is coherent with the 'market is attractive for competitors') or contradictory (i.e., 'our main competitor failed to introduce a competing product' is incoherent with 'new competitors might emerge in our field'). Nodes represent either hypotheses or facts. The activation of each node has a numeric value. Like a system of interconnected springs, the network adjusts the activation of the nodes until it reaches a stable state, satisfying the constraints imposed by the explanatory and contradictory links (see Thagard (1989)).

5.1.6 Calculative cultures: underlying cognitive architectures

Decision-making in risk assessment involves the gathering of information and the identification of causal relationships, thereby deducing a risk assessment from available information (e.g., Fenton *et al.* (2020)). While both Bayesian networks and explanatory coherence have been proposed as the appropriate cognitive architectures to model causal

[53] For a definition of the seven basic principles of ECHO (i.e., symmetry, explanation, analogy, data priority, contradiction, competition, and acceptance) see Thagard (1989).

inference (Thagard, 2004), both are, in principle, suited for this risk assessment task. We argue that the choice depends on the predominant calculative culture to be represented.

Bayesian networks require accurate probabilistic information and calculate a precise output (e.g., a specific value for a risk probability or impact). Such a model is akin to quantitative enthusiasm: every piece of information is quantifiable, and the outcome can be used for further calculations.[54]

ECHO networks, by contrast, cannot incorporate precise values for probabilities. Instead, they rely on qualitative descriptions of relationships. They model how individuals make sense of information and account for the limited capability of individuals to make precise calculations or complex logical deductions. Thus, ECHO networks are well suited to model discussions in a quantitative skepticism setting. In line with this conceptual distinction, we compare the results of the simulation of risk workshops generated by Harten *et al.* (2022) using Bayesian networks to model a quantitative enthusiasm setting with the results of the simulation of risk workshops using ECHO networks to model a quantitative skepticism setting.

5.2 Methods

This section describes the method used for the simulation study. We recap the design of the previous study and highlight the necessary changes needed to compare the impact of different calculative cultures.

5.2.1 Overall design

Like Harten *et al.* (2022), this study uses a simulation experiment approach which combines a model of the processes we want to investigate with an appropriate experimental design (Harrison *et al.*, 2007). The code of the simulation and the ODD+D (Overview, Design Concepts and Details + Decision) protocol, containing a standardized description of the technical implementation of the simulation (Grimm *et al.*, 2006, 2020; Müller *et al.*, 2013) is reported on CoMSES (Bellora-Bienengräber *et al.*, 2022).

[54] See Neil *et al.* (2019) for an example of how Bayesian networks can assist communication in a quantitative enthusiasm context.

To better understand the impact of calculative culture on the outcome of risk workshops, we compare the results produced by Harten *et al.* (2022) with those of an implementation using a different cognitive architecture. While we change the cognitive architecture used, we employ the same approach to modeling the interaction of the participants in the risk workshops. The interaction is modeled as an ABM with information exchange between agents (Harten *et al.*, 2022; Lorscheid and Meyer, 2021; Wall and Leitner, 2021).[55]

5.2.2 Model of the discussion process and risk assessment

The discussion process used for this study is identical to the one used by Harten *et al.* (2022). For each simulation experiment, we conduct multiple simulation runs. Each run is a discussion of a single risk in a risk workshop. The simulation run consists of five stages (Figure 16).

[55] See Davies *et al.* (2010) for a discussion of possible applications of ABM to model risk regulation .

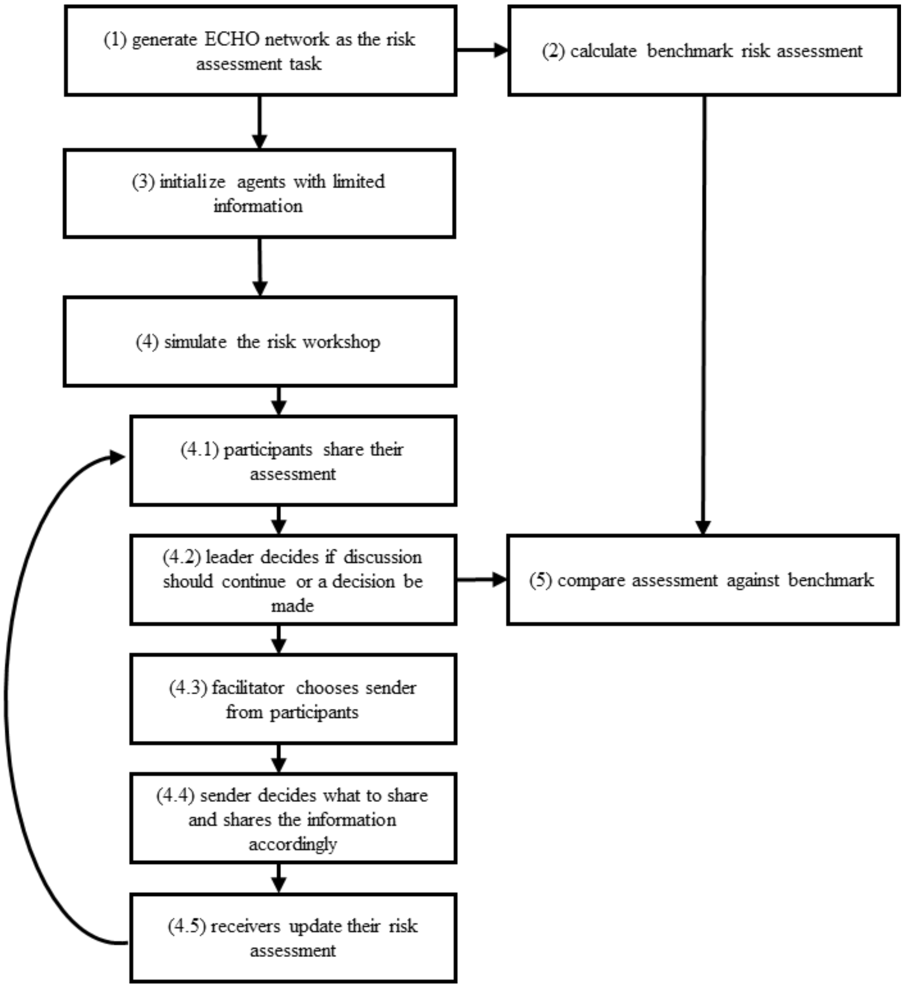


Figure 16 Stages of the simulation before, during, and after the risk workshop.[56]

In the ECHO networks used for this study, we adapt the structure of the Bayesian networks used by Harten *et al.* (2022). Figure 17 compares the structure of the Bayesian network used by Harten *et al.* (2022) with the ECHO networks used in this study. For each discussion, a new risk is generated. The network representing full knowledge of a risk contains 38 nodes, comprising 27 information nodes, nine issue nodes, and two nodes from which the overall risk assessment regarding the likelihood of the risk is derived (see Davies *et al.* (2010) for a similar conceptualization in the context of decision making in risk regulation). Due to differences in their backgrounds or priorities, individual participants start with diverse risk perceptions (Sjöberg, 2000). Initially, they are only aware of

[56] Adapted from Harten *et al.* (2022).

the information and issues they are provided with before the start of the discussion. Information on the 36 information and issue nodes is exchanged during the discussion. If participants hear about a node that they have previously been unaware of, they include it in their mental model. All nodes have an activation between -1 and 1, representing their degree of belief in the underlying information or issue. Depending on their knowledge, participants can reach different risk assessments for the same risk.

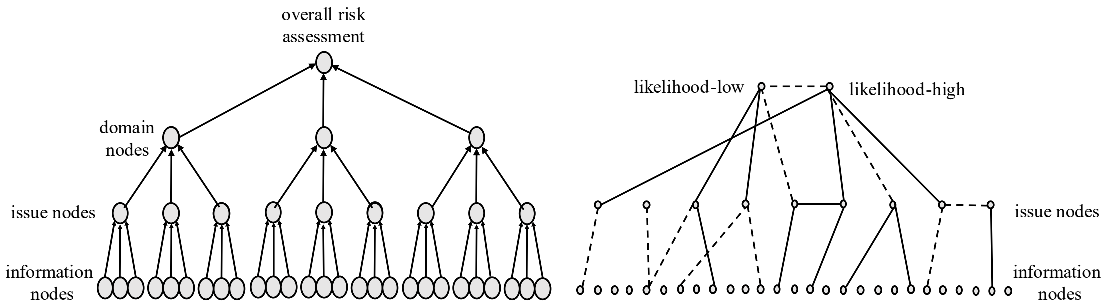


Figure 17 The structure of the Bayesian network used by Harten *et al.* (2022) (left) and the stylized ECHO network used to model quantitative skepticism (right).[57]

5.2.3 Model of the discussion

Like Harten *et al.* (2022), nine participants exchange information in the risk workshop. These participants make their best effort to gain a correct understanding of the risk, i.e., they do not follow their own and potentially hidden agendas. For example, there are no agency conflicts between participants at different hierarchical levels. The discussion consists of discussion rounds, each following a set sequence of participant activities (see

[57] Note: For the ECHO network, full lines indicate explanatory relationships and dotted lines contradictory relationships. For visual clarity, only a selection of the existing relationships is depicted. Relationships can exist between information and issue nodes, among the issue nodes, and between issue nodes and the likelihood nodes used for deriving the risk assessment.

stage 4 in Figure 16). The impact of the distinctive design characteristics of the risk workshop is analyzed in four simulation experiments.

Table 4 Overview of the simulation experiments.[58]

	Simulation experiments			
	Limits to information transfer	Incomplete discussions	Group characteristics	Interaction patterns
Experimental Conditions				
Receivers retain a part of their prior beliefs	**yes**	yes	yes	yes
Leader's decision approach	leader follows own opinion, majority opinion	**leader follows own opinion, consensus, majority opinion**	leader follows majority opinion	leader follows majority opinion
Termination of the discussion	n/a	**at first consensus, after one, five, or ten stable rounds, full discussion**	full discussion	full discussion
Differences in the distribution of information	no	no	**yes/no**	no
Receivers consider hierarchical differences	no	no	**yes/no**	no
Receivers have no transactive memory	no	no	**yes/no**	no
Interaction pattern	random	random	random	**priority to concern vs. dissent vs. hierarchy vs. homogeneity**
Outcome Variables	% of correct assessments per discussion round	% of correct assessments, avg. number of discussion rounds	% of correct assessments, avg. number of discussion rounds	% of correct assessments, avg. number of discussion rounds
Number of simulated discussions (n) [a]	1.129	1.129	11.972	11.463
Number of high / low risks	628 / 601	628 / 601	5.990 / 5982	5.701 / 5.762

Note: [a] Each discussion was simulated until the last piece of information was shared at least once. When deciding on the number of simulation runs, it is necessary to balance computational costs and obtaining representative data generated by the simulation's stochastic process (Lorscheid *et al.*, 2012).

[58] Adapted from Harten et al. (2022).

5.2.4 Simulation experiment 1: limits to information transfer

Receivers of information do not simply adopt the view of the sender of the information as to their own but make sense of the information and integrate it into their respective mental models. Therefore, even if all information is shared during the risk workshop, participants will not necessarily have identical mental models at the end of the workshop. Simulating the information exchange in the risk workshop always allows us to investigate the mental state of all participants and thereby investigate the evolution of the risk assessment of the individuals and, accordingly, of the group.

5.2.5 Simulation experiment 2: incomplete discussions

Every discussion must stop after the elapse of a certain amount of time. In the simulated risk workshop, the decision to end the discussion and decide on a risk assessment is made by the leader (i.e., a designated participant). The leader uses heuristics to determine whether the discussion should end. Harten *et al.* (2022) investigate two heuristics: either the leader requires a consensus of all participants to end the discussion, or the discussion ends after it has stagnated for some time. Here, stagnation is defined as no change in the average (numerical) risk assessment over several rounds.

While Harten *et al.* (2022) use stagnation in the discussion as an indicator to terminate it, this is not suitably applicable to the quantitative skepticism model. In the quantitative enthusiasm model, the assessments tend to converge on the final risk assessment, and a slow rate of change denotes reaching the final stage of the discussion. The same does not apply to the quantitative skepticism model. Here assessment changes happen suddenly, even after long periods of unchanged risk assessments. We, therefore, also simulate another heuristic, i.e., the continuation of the discussion until each piece of information has been mentioned at least once.

5.2.6 Simulation experiment 3: specific group characteristics

We model the effect of the same three group characteristics analyzed by Harten *et al.* (2022), thus accounting for the following group characteristics:

- Differences in the distribution of information. Information is distributed uniformly among the participants before the start of the discussion, or some participants are initially provided with more information than others.
- Differences in hierarchy. In the baseline scenario, participants disregard information shared by someone higher up or lower down in the hierarchy. Alternatively, they can weigh the provided information to reflect the higher belief attached to the information by someone higher up in the hierarchy.
- Information about each other's field of expertise (transactive memory). When participants have a transactive memory, they can give greater weight to information provided by experts. Otherwise, they can dismiss the expertise when processing the information provided.

5.2.7 Simulation experiment 4: specific interaction patterns

Like Harten *et al.* (2022), we simulated five patterns according to which the risk workshop facilitator arranges the order of speakers during the workshop. The baseline scenario assumes that the next speaker is chosen at random. In the remaining patterns, the facilitator selects participants who are concerned about the risk (i.e., participants who assess the risk higher than other participants); participants whose assessment differs the most from the average assessment of the other participants (i.e., dissenters); participants whose assessment is often close to the average assessment of the other participants (i.e., when homogeneity prevails); and, lastly, the facilitator prioritizes participants based on their higher hierarchical position.

5.2.8 Benchmarking the risk workshop

As a baseline for evaluating the effectiveness of a risk assessment workshop, it is necessary to define a benchmark risk assessment. Harten *et al.* (2022) propose a simulated risk workshop under ideal conditions as the benchmark risk assessment. This choice rests on the assumption that performing a discussion using settings of an ideal speech situation—

where experts share only correct information, and all other participants integrate the correct information into their mental model—leads each participant to the correct risk assessment. Indeed, for their quantitative enthusiasm model, Harten *et al.* (2022) show this effect in their benchmark simulation.

In the quantitative skepticism model, we find that the order in which participants learn new information strongly influences the risk assessment they reach, both individually and collectively. Furthermore, depending on their current mental model (i.e., their individual ECHO model when they learn new information), the participants react differently to new information. As there is no clear-cut, ideal order of receiving new information, there is also no mechanism that always grants all participants the collective ability to reach an identical assessment that could serve as a benchmark. Consciously deviating from Harten *et al.* (2022), we, therefore, use the assessment of a hypothetical agent that has all information (i.e., the complete ECHO network representing the risk) from the outset, without any stepwise learning through the risk workshop discussion, as a benchmark risk assessment. This state is analogous to the consensus assessment reached under ideal conditions in the quantitative enthusiasm model of Harten *et al.* (2022) as, for our model, having all information at the outset is conceptually equivalent to the result of learning all information under ideal discourse conditions.

5.3 Results and discussion

The simulation study allows us to investigate the dynamics of the discussion on the individual level and on the collective level. In the following, we present and discuss the results.

5.3.1 Individual dynamics of the discussion

Figure 15 depicts two examples of the evolution of the individual risk assessment of the nine participants during a simulated risk workshop. Although, in both examples, the participants simultaneously receive the same information during the discussion, they reach different conclusions. These assessments are based on the previous state of their respective mental models and can differ accordingly (e.g., information might shift the assessment of one participant but be absorbed by another participant's mental model that is consistent with the new information). During certain risk workshops like the one depicted on the left-hand side of Figure 18, all participants will reach the same overall risk assessment. In other workshops, like the one depicted on the right-hand side of Figure 18, participants split into distinct groups, even after sharing all information.[59]

Given this strong dependency of the risk assessment on the risk workshop participant's previous mental model (and thus on the timing of receiving the information), we cannot expect all participants to reach a consensual risk assessment under these conditions. Thus, to define a benchmark risk assessment, we use the risk assessment made by an agent having all information from the outset (see the *Methods* section).

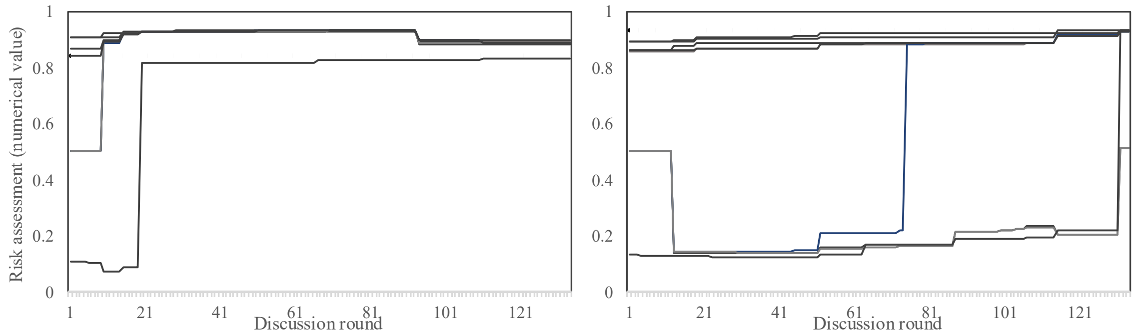


Figure 18 Two examples of individual risk assessments of all nine participants during the risk workshop. [60]

5.3.2 Collective dynamics of the discussion

In the following, we discuss the results of the four simulation experiments concerning the collective dynamics of the simulated risk workshops. The simulation experiments follow the same structure as the experiments presented in Chapter 4.

[59] This pattern is frequently observed in models of opinion dynamics, e.g., Hegselmann and Krause (2002).

[60] The dotted line marks the benchmark risk assessment. The left-hand panel shows a risk workshop where participants reach an assessment close to the benchmark, early in the discussion. The right-hand panel shows a risk workshop where participants change their assessment late in the discussion.

5.3.2.1 Simulation experiment 1: limits to information transfer

As shown in the previous section, under a quantitative skepticism culture, changes in individual risk workshops occur as dramatic shifts in single opinions (see the upward and downward movements in the risk assessment in Figure 18). However, in the aggregate, over many risk workshops, we find that correctness increases continuously. This is clear in Figure 19, which depicts the share of risk workshops that would achieve a correct risk assessment after a certain number of discussion rounds when the leaders follow either their own opinion or the assessment of the majority of the participants. Accordingly, we find that both the leader's own assessment and the majority assessment, on average, continuously improve the correctness of all simulated risk workshops until reaching an upper limit.[61] In comparison, under the condition of quantitative enthusiasm (Harten et al., 2022), individual discussions evolve more gradually. As a result, the aggregated correctness over many risk workshops nevertheless increases faster at the beginning of the risk workshop, with diminishing correctness returns over time. Hence, when a quantitative skepticism culture prevails, it is reasonable—from the viewpoint of improving the correctness of the resulting risk assessment—to continue the discussion until all information has been shared (i.e., after 118 ± 30 discussion rounds; not tabulated). Still, even after such a lengthy discussion, a correct risk assessment is not always achieved. However, when quantitative enthusiasm culture dominates and the participants' risk assessment evolves more gradually, a few stable rounds (after which all information has not yet been shared) may be sufficient to correctly classify a large number of risks.

[61] Before the start of the risk workshop, participants can only rely on the information they are initially provided with. Without any information, a correct assessment—as a purely random choice between two assessments, i.e., 'high' and 'low'—is expected for 50% of the risks. However, the pieces of information provided before the start of the discussion (i.e., in discussion round 0) allow participants to make a correct risk assessment for about 69% of all risks.

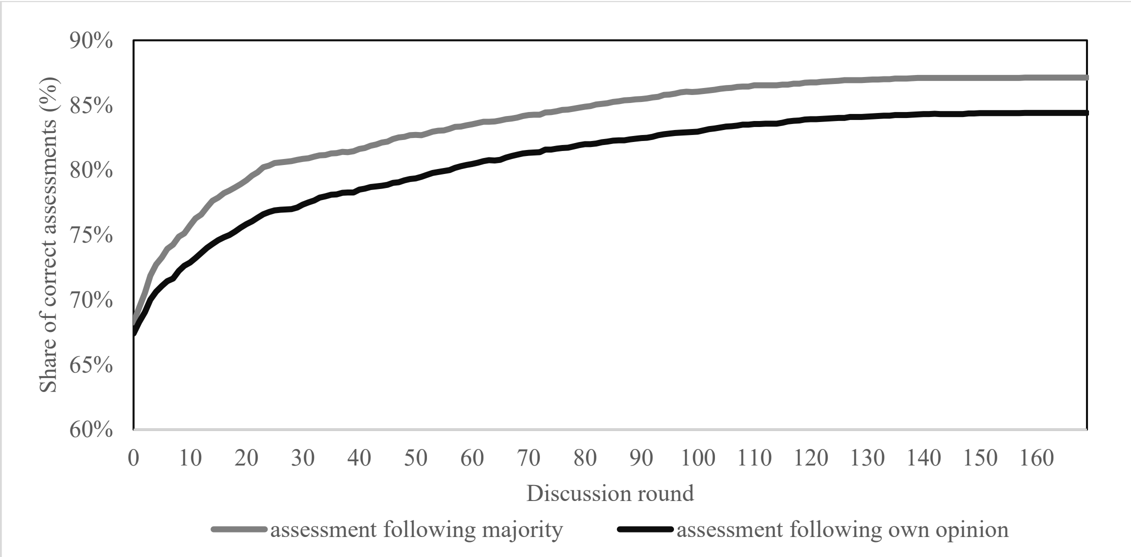


Figure 19 Development of the correctness of the risk assessment over discussion rounds.[62]

.

5.3.2.2 Simulation experiment 2: incomplete discussions

In our model, we find that 60% of discussions (not tabulated) reach a consensus and that this happens on average after 33.4 discussion rounds (Table 5). In these cases, the consensus is correct for 98.1% of all risks. If the discussion is ended based on its stagnation, we find that following the group majority has a slight but consistent advantage across all risks, compared to merely following the leader's individual assessment, independent from the required length of stagnation (either one, five, or ten discussion rounds without a change in the average risk assessment).

If the discussion is only terminated when every piece of information has been discussed at least once, 87% of all risks are correctly assessed, though on average, only after 118 discussion rounds.[63] When comparing the results after one stable round with the results

[62] Note: The figure depicts the share of risk workshops that would reach a correct risk assessment if ended after the current discussion round, following either the majority opinion or the leader's own

[63] The number of discussion rounds is very high compared to the numbers reported by Harten *et al.* (2022), because we focus our investigation on risk workshops that only finish after all

after ten stable rounds, it again becomes apparent that, in a quantitative skepticism culture, a temporarily stagnant discussion is not a good indicator that the risk workshop has reached its best achievable result (see the results above). Due to the comparatively dramatic shifts in opinions particular to this calculative culture, it is necessary to continue the discussion through these phases of (apparent) stagnation to reach the best possible results.

Harten *et al.* (2022) identify a pattern where initially learning about the risk structure increases uncertainty. This mechanism is not present in the quantitative skepticism model, as learning about new information does not introduce uncertainty.[64] New information will be processed and at once converted into a state of coherence with previously available information. What is important in this model is the integration of the information into the mental model and not so much the extent to which the information is believable. For example, the idea that a new competitor might enter the market has a greater impact than the actual probability thereof. However, like Harten *et al.* (2022), we find that participants can also be 'right for the wrong reason,' as they also have a slight tendency to initially overestimate risks. Participants with too little information to inform their risk assessment will default to a 'high' assessment. Subsequently, they will adjust their risk assessment downwards as they learn more, thereby reducing the rate of correctly assessed high risks and increasing the rate of correctly assessed low risks. Herein lies a trade-off between not identifying high risks and overestimating too many low risks.

information has been shared and not, as Harten *et al.* (2022), on risk workshops that finish after ten stable rounds.

[64] In a quantitative enthusiasm model, learning the latest information occurs at two levels. At the first level, participants learn about the existence of a certain node, but not yet about its related probability (i.e., they learn what they do not know). At the second level, they assign a non-zero probability to the possibility that the node is in a 'high risk' state and their uncertainty increases. In later steps, participants attach a probability to the node. The learning at two levels does not occur in a quantitative skepticism model, where learning about the latest information does not introduce uncertainty.

Table 5 Effects of incomplete discussions on risk assessment effectiveness.

Decision-making rule	The proportion of correct assessments			Avg. number of discussion rounds
	All risks (%)	High risks (%)	Low risks (%)	
Stop at first group consensus	98.1%	98.0%	98.2%	33.4
One stable round				
Leader follows own opinion	68.8%	87.1%	49.7%	2.1
Leader follows majority	70.5%	95.8%	43.9%	
Five stable rounds				
Leader follows own opinion	71.8%	84.4%	58.6%	7.6
Leader follows majority	74.7%	91.6%	57.1%	
Ten stable rounds				
Leader follows own opinion	73.2%	83.3%	62.7%	15.5
Leader follows majority	78.3%	88.5%	67.6%	
Full discussion				
Leader follows own opinion	83.0%	85.3%	80.6%	117.8
Leader follows majority	87.0%	86.3%	87.7%	

Note: The table shows the percentage of risks that are correctly assessed and the average number of rounds before the decision is made, depending on the mechanism used to end the discussion. A discussion is said to have *n* stable rounds if the average (numerical) risk assessment does not change more than 2% over *n* consecutive rounds.

5.3.2.3 Simulation experiment 3: group characteristics

Concerning group characteristics, we observe the highest effectiveness of risk assessments when information is equally distributed among the participants (Table 6).[65] This is in line with the results obtained by Harten *et al.* (2022) in a quantitative enthusiasm setting. We do not find a significant impact (not tabulated) of the presence of transactional memory or of taking hierarchical differences into account. In the quantitative skepticism model, compared to the quantitative enthusiasm model, it is less important how strong the first belief in any particular piece of information is (which could be impacted by the hierarchical position or expertise of the sender), as the belief will immediately be adjusted to a coherent state regarding other information already available to the participant. For example, as soon as the issue that ‘a competitor might enter our market’ is introduced to the mental model, it might affect the participant’s risk assessment, even if the activation—

[65] A table that presents simulation results after 10 stable discussion rounds is provided in appendix A.2.

i.e., the degree of belief—is low because, for instance, the information stems from a peer instead of from a superior.

Table 6 Effects of group characteristics on risk assessment effectiveness.

Differences in the distribution of knowledge	Receivers consider hierarchical differences	Receivers have no transactive memory	The proportion of correct assessments			
			All risks (%)	High risks (%)	Low risks (%)	Avg. number of discussion rounds
+	+	+	83.7%	85.4%	81.9%	136.9
+	+	-	80.2%	80.3%	80.1%	136.1
+	-	+	81.3%	80.7%	81.9%	136.3
+	-	-	81.6%	80.3%	83.0%	135.7
-	+	+	85.7%	85.2%	86.2%	117.3
-	+	-	87.4%	87.2%	87.6%	118.7
-	-	+	86.2%	83.3%	89.2%	118.2
-	-	-	87.4%	87.5%	87.3%	118.1

Note: The table shows the percentage of correctly assessed risks after a full discussion when the leader follows the majority vote and the average number of rounds before the decision is made, depending on group characteristics that might influence the risk workshop's effectiveness. A "+" indicates the presence of the corresponding deviation from an ideal discussion situation. Differences in the distribution of information are implemented by giving some participants a higher probability of receiving information during the initialization. If receivers consider hierarchical differences, they weigh input according to the sender's hierarchical position compared to their own. With no transactive memory, participants do not distinguish between senders who are experts on the information and those who are not.

5.3.2.4 Simulation experiment 4: interaction patterns

Only the interaction pattern favoring concerned participants differs remarkably from the baseline interaction scenario following a random order of participants (Table 7).[66] Favoring concerned participants will improve the assessment of high risks but will decrease the rate of correctly identified low risks. Thus, this interaction pattern introduces a bias toward assessing risks as high. Our results differ from Harten *et al.* (2022) regarding the assessment of low risks with prioritized, concerned participants, as they find an improvement in the assessment of all risks, as opposed to the better assessment of high risks only.

[66] A table that presents simulation results after 10 stable discussion rounds is provided in appendix A.3.

Therefore, prioritizing concerned participants comes at the cost of correctly assessing low risks in a quantitative skepticism model. Given such a rule, this results from the focus on high-risk assessments during the early discussion rounds of the risk workshops.

For rules prioritizing hierarchy and homogeneity, Harten *et al.* (2022) also identify significant changes compared to the baseline scenario. These are less pronounced or even absent in the quantitative skepticism model, as it is less important who first introduces a new piece of information as long as it is introduced. Participants will at once assess the information's consistency with their previous mental model. It is also noteworthy that prioritizing participants close to the group opinion will dramatically increase the number of discussion rounds needed to introduce all information, while in Harten et al. (2022), the discussions are ended earlier due to the termination criteria based on sensible stagnation in this quantitative enthusiasm culture.

Table 7 Effects of interaction patterns on risk assessment effectiveness.

	Proportion of correct assessments			
Who talks next during the discussion?	All risks (%)	High risks (%)	Low risks (%)	Avg. number of discussion rounds
Random choice of participants	87.4%	87.5%	87.3%	118.1
Priority to concerned participants	81.6%	95.1%	69.0%	155.7
Priority to participants with dissenting opinions	83.8%	82.1%	85.5%	128.8
Priority to participants with higher hierarchical position	87.8%	87.5%	88.2%	127.1
Priority to participants close to group opinion	86.7%	84.9%	88.5%	737.9

Note: Percentage of risks that are correctly assessed after a full discussion when the leader follows the majority vote and the average number of rounds before the decision is made, depending on the interaction pattern. Concerned participants are those who assess the risk as particularly high. Dissenting participants and those close to the group opinion are determined by measuring the distance between their risk assessment and the average risk assessment of the group.

5.3.2.5 Summary and comparison of calculative cultures

By repeating the simulation experiments performed by Harten *et al.* (2022) and using a model representing a different calculative culture, we find that the drivers of the effectiveness of risk assessments are partially sensitive to the dominant calculative culture. Table 8 summarizes the results of both simulation studies and thereby highlights similarities and differences. We learn from simulation experiment 1 that, notably for both calculative cultures, risk assessment improves during the discussion. Still, path-dependencies are characteristic of discussion processes given quantitative skepticism, i.e., it does not only matter *what* information is exchanged but also *when*. Moreover, there is a potential trade-off between correctly identifying high and low risks in both calculative cultures, as the rate of correct assessments of certain risks decreases during the workshop.

Simulation experiment 2 shows that, with quantitative skepticism, it is more difficult to find the correct time to end the discussion, as the rate of correct risk assessments continuously improves until all information has been shared; stagnation is not a good indicator that the discussion is over. Simulation experiment 3 shows that, while both calculative cultures show better results with an equal distribution of information within the group, quantitative enthusiasm is not negatively affected by a lack of transactive memory or the presence of hierarchical differences. It only matters that all information is made available. The credibility or hierarchical position of the participant introducing the information is less important than the compatibility of the information with the mental model of the receivers.

Regarding the interaction patterns, simulation experiment 4 provides evidence that prioritizing concerned participants will improve the assessment of high risks for both calculative cultures. However, when quantitative skepticism dominates, there is a trade-off with correctly assessing low risks. Otherwise, settings with quantitative skepticism are less impacted by interaction patterns than settings with quantitative enthusiasm.

Table 8 Comparison of results for the simulation experiments with quantitative enthusiasm and quantitative skepticism.

		Calculative culture	
		Quantitative enthusiasm	Quantitative skepticism
Process		Overall improvement of risk assessments over time	Improvement of risk assessments over time
		A trade-off between the identification of high and low risks	A trade-off between the identification of high and low risks
		Gradual changes in individual and collective assessments	Sudden and unpredictable shifts in individual and collective assessments
Design			
	Incomplete discussions	Stagnation in assessments as an indicator to end the discussion	Changes in assessment happen suddenly. Continue discussions even after lengthy periods of stagnation until all information has been shared.
			Following the group majority always has a slight but significant advantage across all risks
	Group characteristics	Differences in the distribution of information have negative effects	Differences in the distribution of information have significant negative effects
		The absence of transactive memory has negative effects	The absence of transactive memory has no significant negative effects
		Hierarchical differences have negative effects	Hierarchical differences have no significant effects
	Interaction patterns	Prioritizing concerned participants improves the assessment of all risks	Prioritizing concerned participants improves the assessment of high risks but lowers the rate of correctly identified low risks
		Prioritizing dissenting participants improves the assessment of high risks but lowers the rate of correctly identified low risks	Prioritizing dissenting participants lowers the rate of correctly assessed risks
		Significant changes compared to the baseline scenario for prioritizing dissent, hierarchy, and homogeneity	No observed effects for prioritizing hierarchy and homogeneity

Note: The results for quantitative enthusiasm are based on Harten *et al.* (2022), while the results for quantitative skepticism are based on this study. Our simulation experiments were designed to allow a comparison.

5.4 Conclusion on the impact of culture

Calculative cultures affect how risk-related information is processed by individuals and translated into risk assessments. This makes it a critical boundary condition for the effectiveness of the design and implementation choices in risk workshops—a common risk assessment technique in organizations. This study explored drivers of the effectiveness of risk assessments in risk workshops, given a calculative culture of quantitative skepticism, and compared them with the findings of previous research that implicitly assumed a calculative culture of quantitative enthusiasm. Using ABM, we modeled individuals' information processing and judgment formation and represented this calculative culture using ECHO models capturing agents' cognition. This allowed us to extend previous simulation experiments of ABM-rooted risk workshops by consciously incorporating the role of the predominant calculative culture in the modeling.

Our results make three contributions to research and practice. First, given the calculative nature of quantitative skepticism, some distinct effects must be considered when designing and implementing risk workshops. Notwithstanding Mikes' (2009) early recognition of the importance of distinguishing between different calculative cultures, our study is, to the best of our knowledge, the first to show the considerable impact that differences in calculative cultures have on the outcome of risk assessments. We show that the predominant calculative culture of an organization is a critical boundary condition when considering the process of risk workshops over time.

Compared to Harten *et al.'s* (2022) findings in a quantitative skepticism culture, the improvement of the correctness of the risk assessment is anything but a gradual process; we rather document sudden shifts in individual and collective assessments. This result questions whether, given this boundary condition, a stagnating discourse is a good indicator to end a discussion. Instead, risk workshop facilitators should create a setting in which everyone is encouraged to share all their information as soon as possible.

We also follow the call of Katzenbach and Smith (2015) to specify rules of interaction. Unlike Harten *et al.* (2022), we do not observe positive effects when prioritizing dissent or negative effects of hierarchy or homogeneity. Additionally, we are the first to document a possible bias when applying the common rule of thumb to prioritize concerned participants during the risk workshop, as supported by Harten *et al.* (2022). We find that favoring concerned participants improves the assessment of high risks but that it lowers the rate of correctly identified low risks. This result indicates the possible existence of a focusing illusion or an anchoring bias (Kahneman, 2011), also for risk workshops. In contrast to Moreland and Myaskovsky (2000) and what is found in Harten *et al.* (2022), we observe no positive effect of a group member's familiarity with the other members' expertise on group performance. Facilitators should be aware of the predominant calculative culture in the organization when deciding whether they need to limit the involvement of dissenters and superiors. Overall, risk workshop facilitators should gain expertise in recognizing the predominant calculative culture in an organization.

Second, we show that some recommendations by Harten *et al.* (2022) are robust and, therefore, resistant to different calculative cultures. Previous research, assuming quantitative enthusiasm, documented a trade-off between the correct assessment of high and low risks (Harten *et al.*, 2022). This trade-off is also present in a predominant quantitative skepticism culture. Facilitators need to be aware that any decision on the design and implementation of risk workshops must take the organization's risk appetite into account. Increasing the correctness of detecting high risks will come at the cost of overestimating low risks, independent of the organization's positioning along the calculative culture continuum. Moreover, we also find that an equal distribution of information within the group leads to the most effective risk workshops in both calculative cultures.

Third, we contribute conceptually and methodologically to the risk literature by introducing a novel approach that allows a model-based investigation of two distinctive ways risk information is processed and used for risk assessments. So far, the literature has mainly used case studies to identify related models (Mikes, 2009). In addition, literature used surveys to quantify the antecedents and consequences of related types of information usage, such as the diagnostic vs. interactive use of information (Simons, 1990) or the use of accounting practices as a computation tool instead of a learning tool (Burchell *et al.*, 1980). While the fine-grained study of the actual usage of information and related cognitive processes is particularly challenging in case studies and even more so in surveys, it can be made amenable for a detailed analysis through ABM. Our ABM made replicable by our in-depth description in the ODD+D, enables future researchers to model the complex interplay between calculative culture, individual cognitive processes, and the related group-level outcomes. Moreover, the scant prior ABM modeling approaches of risk assessments (Harten et al., 2022) do implement but do not conceptualize the modeling of a specific calculative culture. Therefore, to the best of our knowledge, this is the first study on risk assessment that conceptualizes and investigates the role of calculative cultures for individual cognitive processes in combination with other organizational variables.

Like other studies, this study has limitations that future research could address. First, culture is a complex and rich phenomenon. Although we believe that our model represents key facets, a formal representation of calculative cultures via different cognitive architectures cannot capture all nuances of reality. Future research should draw on the rich qualitative work of Mikes (2009) and others to capture facets not included in our study, like the maturity of the risk function. Second, to allow a comparison of the two calculative cultures identified by Mikes (2009), we followed the experimental design of Harten *et al.* (2022) by focusing on the assessment of a single risk and by making a binary distinction between high and low risks. In addition, this also includes their choice of manipulated variables. Future research should search for additional factors that explain differences in the effectiveness of risk assessments in risk workshops. For example, as suggested by Bromiley *et al.* (2014), this could include participants' motivation, hidden agendas, or changes in the groups' ability to transfer the information as the discussion progresses. As ABM allows the incorporation of different personality types (Davies *et al.*, 2010), researchers should also investigate the effects of different portions of extroverted vs. introverted workshop participants and how this interacts with calculative culture and the other risk workshop design characteristics considered in this study. Third, our model is limited to the context of risk workshops taking place with participants participating synchronously in the discussion. Thus, it does not include risk assessments that happen with dispersed participants interacting asynchronously with each other (e.g., via e-mail). While it is likely that many effects identified in this study would manifest also when considering the latter type of risk assessments, future research is called to broaden the applicability of our model by incorporating, for example the effect of time delays in the communication or of selective attention to certain pieces of information. Fourth, the external validity of our results could be strengthened by future research through empirical tests in the form of field and case studies. It might, for example, be interesting to study a potential clash of different calculative cultures in the same risk workshop or organization. Finally, the use of the approach developed in this study to investigate the effects of different calculative practices in other contexts beyond risk workshops and even beyond risk management, appears to be promising.

6 Conclusion

The previous two chapters present the results of two simulation experiments that aimed to answer two research questions: How to facilitate an effective risk workshop? And how should a risk workshop account for the predominant calculative culture in order to be effective? By modeling risk workshops both on the level of interaction and the level of cognition, we provide a testbed for interaction patterns, group characteristics, and decision-making rules in the context of a risk workshop. By comparing the workshop results to an ideal benchmark, we are able to judge the effectiveness of risk workshops under certain conditions. Furthermore, by implementing two different cognitive architectures that correspond to two calculative cultures, we are able to identify how the impact of these factors depends on the predominant culture. We find that workshop facilitators need to consider the predominant culture as well as the risk appetite of the organization when deciding on how to conduct the workshop. Beyond the design decisions under the direct control of the workshop facilitator, we highlight the impact of the predominant interaction pattern within the group, an effect that a workshop facilitator should be aware of.

Of course, a simulation experiment can usually just contribute one part of answering a research question about a system as complex as a risk workshop, a challenging task performed by a group of experts in face-to-face interaction. As stated earlier, a simulation experiment can just perform experiments on a model of a subject, not the subject itself (Gilbert and Troitzsch, 2005). Thus, the question arises whether the results of the simulation experiment are applicable to actual risk workshops. We have followed best practices in building the model and interpreting the results to ensure that a high degree of internal validity is attained (cf. Davis *et al.*, 2007). Still, it would be worthwhile to challenge our findings, for example with a qualitative observational study of actual risk workshops or by interviewing risk workshop facilitators.

Apart from the question of internal validity, our study only investigates a limited number of risk workshop characteristics that can impact risk workshop effectiveness. In order to determine the importance of culture for all experiments in the first study, we used the same set of risk workshop characteristics in the second study. However, there are other characteristics than those considered that might influence the effectiveness of risk workshops. For example, we only simulate discussions of isolated risks, while actual risk workshops usually discuss several risks, one after another (Quail, 2011). A facilitator needs to decide on the order of risks discussed and allocate a limited time budget for those risks. Investigating this effect would require accounting for the effects of the discussion of a single risk on the overall workshop. Also, we have only considered risks of the same complexity. In actual risk management practice, some risks will need more deliberation and expertise in order to be assessed than others. Again, it would be worthwhile investigating how risk complexity should inform the design of risk workshops. When investigating the importance of culture, we assume that all participants subscribe to the dominant culture within the organization. It would be interesting to investigate the effects of having a group of participants that do not share one calculative culture. Finally, while we use a complex model for the cognition of the participants, we use a comparatively simple model for the interaction of participants. We do not account for deviations regarding the motivation of the participants (e.g., a hidden agenda to force a specific risk assessment or to hide specific information).

One main benefit of simulation experiments is the potential to provide full transparency on how the study was conducted, as well as allow other researchers to build upon the actual simulation models used for the experiments. The code for both simulation studies, written in Python and R, has been published along with documentation under the free GNU General Public License 3 on the public CoMSES repository (Bellora-Bienengräber *et al.*, 2022; Harten *et al.*, 2021). This way, we provide a testbed for risk workshops that can be used to conduct further studies on this topic, including addressing the limitations mentioned above. Also, parts of the code can be reused for simulation studies on other subjects that require complex cognitive architectures for agents in an agent-based simulation. For example, to the best of our knowledge, we provide the first implementation of ECHO written in Python.

The two studies conducted for this thesis simulate a very specific part of the risk management process. This raises the question of what we can learn from these studies for, on the one hand, the broader context of risk management and, on the other hand, for similar group work on other topics. In the following, I discuss the external validity of the study in those two directions.

Within the risk management process, we limit the scope of the risk workshop to a very limited task: The assessment of an already identified risk on a one-dimensional categorical scale. However, risk workshops are often also used for risk identification, and risk assessment often happens in more than one dimension. A model that accounts for both these aspects would be possible with a very similar setup to the one we chose for our studies. Risk identification relies on gathering information distributed among different stakeholders that might indicate that a non-trivial risk is present in a certain area. This information-gathering task is structurally similar to how agents in our model combine their expertise to gain a full understanding of a risk. The approach to dissent within the group regarding a risk would be different, however. During risk identification, an organization can generally err on the side of caution to avoid missing critical risks. During risk assessment, a final judgment is needed for each risk, because risk assessment provides the basis for further actions taken by the organization to address the risks. Thus, when facilitating a risk workshop aimed at risk identification, it would be advisable to follow a process that has a high rate of correctly assessed high risk (at the cost of incorrectly assessed low risks).

While the model only accounts for a one-dimensional risk assessment, the results are equally applicable to a multi-dimensional risk assessment. The same model could be used to assess several dimensions of the same risk, only the model of the risk itself would need to be appended to include information regarding all dimensions under investigation. The choice to limit the model to a single dimension was made to reduce the complexity of the result analysis.

While many parts of the risk management process happen in collaboration with multiple stakeholders, this collaboration is not always conducted in a face-to-face setting. Often, risk management processes happen asynchronously, using digital tools. Because our models are built upon the idea of a free exchange of information during an open discussion, the results have only limited validity for practices that happen outside of a workshop setting. For example, if multiple stakeholders asynchronously collaborate by filling in a risk register, they might not share their reasoning to the same degree as during an open discussion. While distributed cognition does not require direct exchange and can happen over a medium (cf. Hutchins, 2000), the model would need to reflect the different patterns of exchange.

The model also has some validity for group work outside the domain of risk management. The model of interaction we use is derived from the idea of an ideal speech situation as described by Habermas (1982). Habermas' ideal speech situation is meant to describe ideal conditions for each discussion where participants are interested in finding a consensus, as needed, for example, to define public policies. The model of interaction is, therefore, applicable to any setting where participants work together with an open mind to find the best possible solution for a problem. In the context of an organization, this could also be a strategy workshop or a creative workshop to make product or marketing decisions. In those cases, however, a different model of the subject of the discussion would be required. Both the Bayesian network and the constraint satisfaction network used to model cognition regarding a risk are specific to the typical structure of information informing a risk-related decision. While Bayesian networks and constraint satisfaction networks might still be valid cognitive architectures for these settings (Darwiche, 2009; Thagard, 2000), the decision problem would be structured differently.

Overall, this thesis addresses the need for research into the effectiveness of risk management practices (Aven, 2012; Bromiley *et al.*, 2014). It provides practical guidance for risk management practitioners on how to conduct effective risk workshops, depending on the risk appetite of the organization and the predominant calculative culture. Furthermore, it provides an experimental testbed that can be used for further research on risk management practices as well as group work in general. It also contributes to the research field of agent-based modeling of social systems by advancing the use of complex cognitive architectures to model agent cognition.

7 References

Ackermann, F., Howick, S., Quigley, J., Walls, L. and Houghton, T. (2014), "Systemic risk elicitation: using causal maps to engage stakeholders and build a comprehensive view of risks", *European Journal of Operational Research*, Vol. 238 No. 1, pp. 290–299, doi: 10.1016/j.ejor.2014.03.035.

Anderson, J.R. (1983), *The Architecture of Cognition*, Harvard University Press, Cambridge, Mass.

Arbis, D., Dixit, V.V. and Rashidi, T.H. (2016), "Impact of risk attitudes and perception on game theoretic driving interactions and safety", *Accident Analysis & Prevention*, Vol. 94, pp. 135–142, doi: 10.1016/j.aap.2016.05.027.

Arena, M., Arnaboldi, M. and Azzone, G. (2011), "Is enterprise risk management real?", *Journal of Risk Research*, Routledge, Vol. 14 No. 7, pp. 779–797, doi: 10.1080/13669877.2011.571775.

van Asselt, M.B.A. and Renn, O. (2011), "Risk governance", *Journal of Risk Research*, Vol. 14 No. 4, pp. 431–449, doi: 10.1080/13669877.2011.553730.

Aven, T. (2012), "Foundational Issues in Risk Assessment and Risk Management", *Risk Analysis*, Vol. 32 No. 10, pp. 1647–1656, doi: 10/f4bsk8.

Aven, T. (2019), *The Science of Risk Analysis: Foundation and Practice*, 1st ed., Routledge, doi: 10.4324/9780429029189.

Aven, T., Ben-Haim, Y., Andersen, H.B., Cox, T., Droguett, E.L., Greenberg, M., Guikema, S., *et al.* (n.d.). "Society for Risk Analysis Glossary", p. 9.

Aven, T. and Renn, O. (2010), *Risk Management and Governance: Concepts, Guidelines and Applications*, Springer, Heidelberg ; New York.

Aven, T. and Renn, O. (2020), “Some foundational issues related to risk governance and different types of risks”, *Journal of Risk Research*, Vol. 23 No. 9, pp. 1121–1134, doi: 10.1080/13669877.2019.1569099.

Aven, T. and Zio, E. (2014), “Foundational issues in risk assessment and risk management”, *Risk Analysis*, Vol. 34 No. 7, pp. 1164–1172, doi: https://doi.org/10.1111/risa.12132.

Axelrod, R. (1997), “Advancing the Art of Simulation in the Social Sciences”, in Conte, R., Hegselmann, R. and Terna, P. (Eds.), *Simulating Social Phenomena*, Springer, Berlin, Heidelberg, pp. 21–40, doi: 10.1007/978-3-662-03366-1_2.

Bajtelsmit, V., Coats, J.C. and Thistle, P. (2015), “The effect of ambiguity on risk management choices: An experimental study”, *Journal of Risk and Uncertainty*, Vol. 50 No. 3, pp. 249–280, doi: 10.1007/s11166-015-9218-3.

Bao, C., Wu, D., Wan, J., Li, J. and Chen, J. (2017), “Comparison of Different Methods to Design Risk Matrices from the Perspective of Applicability”, *Procedia Computer Science*, Elsevier B.V., Vol. 122, pp. 455–462, doi: 10.1016/j.procs.2017.11.393.

Bardone, E. and Secchi, D. (2017), “Inquisitiveness: distributing rational thinking”, *Team Performance Management: An International Journal*, Emerald Publishing Limited, Vol. 23 No. 1/2, pp. 66–81, doi: 10.1108/TPM-10-2015-0044.

Barth, R., Meyer, M. and Spitzner, J. (2011), “Typical Pitfalls of Simulation Modeling - Lessons Learned from Armed Forces and Business”, *Journal of Artificial Societies and Social Simulation*, Vol. 15 No. 2, p. 5.

Beck, M. and Kewell, B. (2014), *Risk: A Study of Its Origins, History and Politics*, World Scientific, New Jersey.

Beck, U. (1992), *Risk Society: Towards a New Modernity*, Sage Publications, London ; Newbury Park, Calif.

Bellora-Bienengräber, L., Harten, C. and Meyer, M. (2022), "Agent-based simulation of discussion processes in risk workshops with quantitative skepticism v1.0.0", *CoMSES Computational Model Library*, CoMSES.Net.

Bellora-Bienengräber, L., Harten, C. and Meyer, M. (2023), "The effectiveness of risk assessments in risk workshops: the role of calculative cultures", *Journal of Risk Research*, Routledge, Vol. 26 No. 2, pp. 163–183, doi: 10.1080/13669877.2022.2108120.

Boholm, Å. and Corvellec, H. (2016), "The role of valuation practices for risk identification", in Power, M. (Ed.), *Riskwork*, Oxford University Press, Oxford, pp. 110–129, doi: 10.1093/acprof:oso/9780198753223.003.0006.

Borysov, S.S., Rich, J. and Pereira, F.C. (2019), "How to generate micro-agents? A deep generative modeling approach to population synthesis", *TRANSPORTATION RESEARCH PART C-EMERGING TECHNOLOGIES*, Vol. 106, pp. 73–97, doi: 10.1016/j.trc.2019.07.006.

Bromiley, P., McShane, M., Nair, A. and Rustambekov, E. (2014), "Enterprise risk management: review, critique, and research directions", *Long Range Planning*, Vol. 48 No. 4, pp. 265–276, doi: 10.1016/j.lrp.2014.07.005.

Burchell, S., Clubb, C., Hopwood, A., Hughes, J. and Nahapiet, J. (1980), "The roles of accounting in organizations and society", *Accounting, Organizations and Society*, Vol. 5 No. 1, pp. 5–27, doi: 10.1016/0361-3682(80)90017-3.

Callahan, C. and Soileau, J. (2017), "Does Enterprise risk management enhance operating performance?", *Advances in Accounting*, Vol. 37, pp. 122–139, doi: 10.1016/j.adiac.2017.01.001.

Chapman, R.J. (1998), “The effectiveness of working group risk identification and assessment techniques”, *International Journal of Project Management*, Vol. 16 No. 6, pp. 333–343, doi: 10.1016/S0263-7863(98)00015-5.

Cooke, R.M. and Goossens, L.L.H.J. (2008), “TU Delft expert judgment data base”, *Reliability Engineering & System Safety*, Vol. 93 No. 5, pp. 657–674, doi: 10.1016/j.ress.2007.03.005.

COSO. (2004), *Enterprise Risk Management: Integrated Framework*.

COSO. (2017), *Enterprise Risk Management: Aligning Risk with Strategy and Performance*, COSO New York, NY.

Cuevas, E. (2020), “An agent-based model to evaluate the COVID-19 transmission risks in facilities”, *Computers in Biology and Medicine*, Vol. 121, p. 103827, doi: 10.1016/j.compbiomed.2020.103827.

Dam, K.H.V., Bustos-turu, G. and Shah, N. (2011), “A methodology for simulating synthetic populations for the analysis of socio-technical infrastructures”.

Darwiche, A. (2009), *Modeling and Reasoning with Bayesian Networks*, Cambridge University Press, Cambridge, doi: 10.1017/CBO9780511811357.

Davies, G.J., Kendall, G., Soane, E., Li, J., Charnley, F. and Pollard, S.J.T. (2010), “Regulators as ‘agents’: power and personality in risk regulation and a role for agent-based simulation”, *Journal of Risk Research*, Routledge, Vol. 13 No. 8, pp. 961–982, doi: 10.1080/13669877.2010.486075.

Davis, J.P., Eisenhardt, K.M. and Bingham, C.B. (2007), “Developing Theory Through Simulation Methods”, *Academy of Management Review*, Vol. 32 No. 2, pp. 480–499, doi: 10.5465/amr.2007.24351453.

Dawson, R.J., Peppe, R. and Wang, M. (2011), “An agent-based model for risk-based flood incident management”, *Natural Hazards*, Vol. 59 No. 1, pp. 167–189, doi: 10.1007/s11069-011-9745-4.

Dechter, R. (2013), “Introduction”, in Dechter, R. (Ed.), *Reasoning with Probabilistic and Deterministic Graphical Models: Exact Algorithms*, Springer International Publishing, Cham, pp. 1–8, doi: 10.1007/978-3-031-01566-3_1.

Edmonds, B. and Moss, S. (2005), “From KISS to KIDS – An ‘Anti-simplistic’ Modelling Approach”, in Davidsson, P., Logan, B. and Takadama, K. (Eds.), *Multi-Agent and Multi-Agent-Based Simulation*, Springer, Berlin, Heidelberg, pp. 130–144, doi: 10.1007/978-3-540-32243-6_11.

Ekmekci, O. and Casey, A. (2011), “Computer simulation exploring organizational identification for contingent workers”, edited by Lemmergaard, J.*Team Performance Management: An International Journal*, Emerald Group Publishing Limited, Vol. 17 No. 5/6, pp. 279–298, doi: 10.1108/13527591111159018.

Farrell, M. and Gallagher, R. (2015), “The Valuation Implications of Enterprise Risk Management Maturity: V ALUATION I MPLICATIONS OF ERM”, *Journal of Risk and Insurance*, Vol. 82 No. 3, pp. 625–657, doi: 10.1111/jori.12035.

Fenton, N.E. and Neil, M. (2019), *Risk Assessment and Decision Analysis with Bayesian Networks*, Second edition., CRC Press, Taylor & Francis Group, Boca Raton.

Fenton, N.E., Neil, M., Osman, M. and McLachlan, S. (2020), “COVID-19 infection and death rates: the need to incorporate causal explanations for the data and avoid bias in testing”, *Journal of Risk Research*, pp. 1–4, doi: 10.1080/13669877.2020.1756381.

Ferdous, R., Khan, F., Sadiq, R., Amyotte, P. and Veitch, B. (2012), “Handling and updating uncertain information in bow-tie analysis”, *Journal of Loss Prevention in the Process Industries*, Vol. 25 No. 1, pp. 8–19, doi: 10.1016/j.jlp.2011.06.018.

Fioretti, G. (2013), “Agent-Based Simulation Models in Organization Science”, *Organizational Research Methods*, Vol. 16 No. 2, pp. 227–242, doi: 10.1177/1094428112470006.

Fraser, J.R.S. and Simkins, B.J. (2016), “The challenges of and solutions for implementing enterprise risk management”, *Business Horizons*, Vol. 59 No. 6, pp. 689–698, doi: 10.1016/j.bushor.2016.06.007.

French, S., Hanea, A.M., Bedford, T. and Nane, G.F. (2021), “Introduction and Overview of Structured Expert Judgement”, in Hanea, A.M., Nane, G.F., Bedford, T. and French, S. (Eds.), *Expert Judgement in Risk and Decision Analysis*, Springer International Publishing, Cham, pp. 1–16, doi: 10.1007/978-3-030-46474-5_1.

Freuder, E.C. and Wallace, R.J. (1992), “Partial constraint satisfaction”, *Artificial Intelligence*, Vol. 58 No. 1–3, pp. 21–70, doi: 10.1016/0004-3702(92)90004-H.

Frigotto, M.L. and Rossi, A. (2015), “An explanatory coherence model of decision making in ill-structured problems”, *Mind & Society*, Vol. 14 No. 1, pp. 35–55, doi: 10.1007/s11299-014-0158-4.

Giddens, A. (1999), “Reith Lecture 2: Risk”, *London: BBC.*

Gilbert, G.N. and Troitzsch, K.G. (2005), *Simulation for the Social Scientist*, 2nd ed., Open University Press, Maidenhead, England ; New York, NY.

Gilbert, N. (2005), “When does social simulation need cognitive models?”, *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, pp. 428–432, doi: 10.1017/CBO9780511610721.020.

Goerlandt, F. and Reniers, G. (2016), "On the assessment of uncertainty in risk diagrams", *Safety Science*, Elsevier Ltd, Vol. 84, pp. 67–77, doi: 10.1016/j.ssci.2015.12.001.

González-Brenes, J.P., Behrens, J.T., Mislevy, R.J., Levy, R. and DiCerbo, K.E. (2016), "Bayesian Networks", in Rupp, A.A. and Leighton, J.P. (Eds.), *The Handbook of Cognition and Assessment*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 328–353, doi: 10.1002/9781118956588.ch14.

Goossens, L.H.J., Cooke, R.M., Hale, A.R. and Rodić-Wiersma, Lj. (2008), "Fifteen years of expert judgement at TUDelft", *Safety Science*, Vol. 46 No. 2, pp. 234–244, doi: 10.1016/j.ssci.2007.03.002.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., *et al.* (2006), "A standard protocol for describing individual-based and agent-based models", *Ecological Modelling*, Vol. 198 No. 1–2, pp. 115–126, doi: 10.1016/j.ecolmodel.2006.04.023.

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J. and Railsback, S.F. (2010), "The ODD protocol: A review and first update", *Ecological Modelling*, Elsevier B.V., Vol. 221 No. 23, pp. 2760–2768, doi: 10.1016/j.ecolmodel.2010.08.019.

Grimm, V., Polhill, G. and Touza, J. (2013), "Documenting Social Simulation Models: The ODD Protocol as a Standard", Springer, Cham, pp. 117–133, doi: 10.1007/978-3-540-93813-2_7.

Grimm, V., Railsback, S.F., Vincenot, C.E., Berger, U., Gallagher, C., DeAngelis, D.L., Edmonds, B., *et al.* (2020), "The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism", *Journal of Artificial Societies and Social Simulation*, Vol. 23 No. 2, p. 7.

Gupta, U.G. and Clarke, R.E. (1996), “Theory and applications of the Delphi technique: A bibliography (1975–1994)”, *Technological Forecasting and Social Change*, Vol. 53 No. 2, pp. 185–211, doi: 10.1016/S0040-1625(96)00094-7.

Habermas, J. (1971), “Vorbereitende Bemerkungen zu einer Theorie der kommunikativen Kompetenz”, *Theorie Der Gesellschaft Oder Sozialtechnologie - Was Leistet Die Systemforschung?*, Suhrkamp, Frankfurt am Main.

Habermas, J. (1982), *Theorie des kommunikativen Handelns*, 2nd ed., Suhrkamp, Frankfurt am Main.

Habermas, J. (1983), *Moralbewußtsein und kommunikatives Handeln*, Vol. 422, Suhrkamp, Frankfurt am Main.

Habermas, J. (1989), *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns*, 3rd ed., Suhrkamp, Frankfurt am Main.

Haig, B.D. (2009), “Inference to the best explanation: A neglected approach to theory appraisal in psychology”, *The American Journal of Psychology*, University of Illinois Press, Vol. 122 No. 2, pp. 219–234.

Handy, C.B. (1986), *Understanding Organizations*, Penguin, Harmondsworth.

Harrison, J.R., Zhiang Lin, Carroll, G.R. and Carley, K.M. (2007), “Simulation modeling in organizational and management research”, *Academy of Management Review*, Vol. 32 No. 4, pp. 1229–1245, doi: 10.5465/AMR.2007.26586485.

Harten, C. (2019), “Agent-Based Model of Risk Assessment: A Distributed Cognition Approach”, in Linsley, P., Shrives, P. and Wieczorek-Kosmala, M. (Eds.), *Multiple Perspectives in Risk and Risk Management*, Springer International Publishing, pp. 169–178.

Harten, C., Meyer, M. and Bellora-Bienengräber, L. (2021), "An agent-based simulation of discussion processes in risk workshops v1.0.2", *CoMSES Computational Model Library*, CoMSES.Net, 30 September, available at: https://www.comses.net/codebases/424b4103-5cf3-4552-9928-615e75a27ce7/releases/1.0.2/ (accessed 15 June 2022).

Harten, C., Meyer, M. and Bellora-Bienengräber, L. (2022), "Talking about the likelihood of risks: an agent-based simulation of discussion processes in risk workshops", *Journal of Accounting & Organizational Change*, Vol. 18 No. 1, pp. 153–173, doi: 10.1108/JAOC-11-2020-0197.

Hauke, J., Lorscheid, I. and Meyer, M. (2018), "Individuals and their interactions in demand planning processes: an agent-based, computational testbed", *International Journal of Production Research*, Vol. 56 No. 13, pp. 4644–4658, doi: 10.1080/00207543.2017.1377356.

He, H., Martinsson, P. and Sutter, M. (2012), "Group decision making under risk: an experiment with student couples", *Economics Letters*, Vol. 117 No. 3, pp. 691–693, doi: 10.1016/j.econlet.2011.12.081.

Heemstra, F.J., Kusters, R.J. and de Man, H. (2003), "Guidelines for managing bias in project risk management", *2003 International Symposium on Empirical Software Engineering, 2003. ISESE 2003. Proceedings.*, presented at the 2003 International Symposium on Empirical Software Engineering, 2003. ISESE 2003. Proceedings., pp. 272–280, doi: 10.1109/ISESE.2003.1237988.

Hegselmann, R. and Krause, U. (2002), "Opinion dynamics and bounded confidence models, analysis, and simulation", *Journal of Artificial Societies and Social Simulation*, Vol. 5 No. 3.

Hiebl, M.R.W., Baule, R., Dutzi, A., Stein, V. and Wiedemann, A. (2018), “Guest editorial”, *The Journal of Risk Finance*, Vol. 19 No. 4, pp. 318–326, doi: 10.1108/JRF-08-2018-194.

Hiltz, S.R., Johnson, K. and Turoff, M. (1986), “Experiments in group decision making communication process and outcome in face-to-face versus computerized conferences”, *Human Communication Research*, Vol. 13 No. 2, pp. 225–252, doi: https://doi.org/10.1111/j.1468-2958.1986.tb00104.x.

Hopkin, P. and Thompson, C. (2022), *Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Enterprise Risk Management*, Sixth Edition., Kogan Page, London ; New York.

Hora, S.C. (1996), “Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management”, *Reliability Engineering & System Safety*, Vol. 54 No. 2–3, pp. 217–223, doi: 10.1016/S0951-8320(96)00077-4.

Horlick-Jones, T., Rosenhead, J., Georgiou, I., Ravetzd, J. and Löfstedte, R. (2001), “Decision support for organisational risk management by problem structuring”, *Health, Risk & Society*, Vol. 3 No. 2, pp. 141–165, doi: 10.1080/13698570125225.

Hunziker, S. (2019), *Enterprise Risk Management: Modern Approaches to Balancing Risk and Reward*, Gabler Verlag, Wiesbaden, doi: 10.1007/978-3-658-25357-8.

Hutchins, E. (2000), *Cognition in the Wild*, MIT Press, Cambridge, Massachusetts.

ISO. (2009a), *ISO 31000:2009 - Risk Management - Principles and Guidelines*.

ISO. (2009b), “IEC 31010:2009. Risk management — Risk assessment techniques”, Vol. 2009, p. 92.

Janis, I.L. (1972), *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*, Houghton, Mifflin, Boston.

Jensen, F.V. and Nielsen, T.D. (2007), *Bayesian Networks and Decision Graphs*, 2nd ed., Springer, New York.

Johnson, M.J. and Pajares, F. (1996), "When shared decision making works: a 3-year longitudinal study", *American Educational Research Journal*, American Educational Research Association, Vol. 33 No. 3, pp. 599–627, doi: 10.3102/00028312033003599.

Kabir, G., Tesfamariam, S., Francisque, A. and Sadiq, R. (2015), "Evaluating risk of water mains failure using a Bayesian belief network model", *European Journal of Operational Research*, Vol. 240 No. 1, pp. 220–234, doi: 10.1016/j.ejor.2014.06.033.

Kabir, S. and Papadopoulos, Y. (2019), "Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: a review", *Safety Science*, Vol. 115, pp. 154–175, doi: 10.1016/j.ssci.2019.02.009.

Kahneman, D. (2011), *Thinking, Fast and Slow*, Straus and Giroux, New York.

Katzenbach, J.R. and Smith, D.K. (2015), *Wisdom of Teams: Creating the High-Performance Organization*, Harvard Business Review Press, Boston, Massachusetts.

Keil, M., Wallace, L., Turk, D., Dixon-Randall, G. and Nulden, U. (2000), "An investigation of risk perception and risk propensity on the decision to continue a software development project", *Journal of Systems and Software*, Vol. 53 No. 2, pp. 145–157, doi: 10.1016/S0164-1212(00)00010-8.

Kelman, H.C. (1958), "Compliance, identification, and internalization three processes of attitude change", *Journal of Conflict Resolution*, SAGE Publications Inc, Vol. 2 No. 1, pp. 51–60, doi: 10.1177/002200275800200106.

Kewell, B. and Linsley, P. (2017), “Risk tools and risk technologies”, in Woods, M. and Linsley, P. (Eds.), *The Routledge Companion to Accounting and Risk*, Routledge, London, doi: 10.4324/9781315716756-3.

Kim, D.-Y. and Park, J. (2010), “Cultural differences in risk: the group facilitation effect”, *Judgment and Decision Making*, Vol. 5 No. 5, p. 11.

van Knippenberg, D., De Dreu, C.K.W. and Homan, A.C. (2004), “Work group diversity and group performance: an integrative model and research agenda”, *Journal of Applied Psychology*, Vol. 89 No. 6, pp. 1008–1022, doi: 10.1037/0021-9010.89.6.1008.

Kocabas, V. and Dragicevic, S. (2013), “Bayesian networks and agent-based modeling approach for urban land-use and population density change: a BNAS model”, *Journal of Geographical Systems*, Vol. 15 No. 4, pp. 403–426, doi: 10.1007/s10109-012-0171-2.

Koski, T. and Noble, J.M. (2009), *Bayesian Networks: An Introduction*, Wiley, Chichester, West Sussex, UK.

Labro, E. and Vanhoucke, M. (2007), “A simulation analysis of interactions among errors in costing systems”, *The Accounting Review*, American Accounting Association, Vol. 82 No. 4, pp. 939–962.

Lake, B.M., Ullman, T.D., Tenenbaum, J.B. and Gershman, S.J. (2017), “Building machines that learn and think like people”, *Behavioral and Brain Sciences*, Vol. 40, doi: 10.1017/S0140525X16001837.

Lam, J. (2013), *Enterprise Risk Management: From Incentives to Controls*, Second edition., Wiley, Hoboken, New Jersey.

Langley, P., Laird, J.E. and Rogers, S. (2009), "Cognitive architectures: Research issues and challenges", *Cognitive Systems Research*, Vol. 10 No. 2, pp. 141–160, doi: 10.1016/j.cogsys.2006.07.004.

Lauritzen, S.L. and Spiegelhalter, D.J. (1988), "Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 50 No. 2, pp. 157–194, doi: 10.1111/j.2517-6161.1988.tb01721.x.

Lee, J.S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., Voinov, A., *et al.* (2015), "The complexities of agent-based modeling output analysis", *JASSS*, doi: 10.18564/jasss.2897.

Lee, J.S., Keil, M. and Shalev, E. (2019), "Seeing the Trees or the Forest? The Effect of IT Project Managers' Mental Construal on IT Project Risk Management Activities", *Information Systems Research*, Vol. 30 No. 3, pp. 1051–1072, doi: 10.1287/isre.2019.0853.

Lewis, K. (2004), "Knowledge and performance in knowledge-worker teams: a longitudinal study of transactive memory systems", *Management Science*, Vol. 50 No. 11, pp. 1519–1533, doi: 10.1287/mnsc.1040.0257.

LiCalzi, M. and Surucu, O. (2012), "The power of diversity over large solution spaces", *Management Science*, Vol. 58 No. 7, pp. 1408–1421, doi: 10.1287/mnsc.1110.1495.

Linstone, H.A. and Turoff, M. (1975), *The Delphi Method*, Addison-Wesley, Reading, MA.

Lorscheid, I., Heine, B.-O. and Meyer, M. (2012), “Opening the ‘black box’ of simulations: increased transparency and effective communication through the systematic design of experiments”, *Computational and Mathematical Organization Theory*, Vol. 18 No. 1, pp. 22–62, doi: 10.1007/s10588-011-9097-3.

Lorscheid, I. and Meyer, M. (2020), “(Ir-)rationality of teams: a process-oriented model of team cognition emergence”, in Verhagen, H., Borit, M., Bravo, G. and Wijermans, N. (Eds.), *Advances in Social Simulation*, Springer International Publishing, Cham, pp. 273–286, doi: 10.1007/978-3-030-34127-5_26.

Lorscheid, I. and Meyer, M. (2021), “Toward a better understanding of team decision processes: combining laboratory experiments with agent-based modeling”, *Journal of Business Economics*, Vol. 91 No. 9, pp. 1431–1467, doi: 10.1007/s11573-021-01052-x.

Lu, L., Yuan, Y.C. and McLeod, P.L. (2012), “Twenty-five years of hidden profiles in group decision making: A meta-analysis”, *Personality and Social Psychology Review*, SAGE Publications Inc, Vol. 16 No. 1, pp. 54–75, doi: 10.1177/1088868311417243.

Luhmann, N. (1993), “Die Moral des Risikos und das Risiko der Moral”, in Bechmann, G. (Ed.), *Risiko und Gesellschaft*, VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 327–338, doi: 10.1007/978-3-322-90741-7_12.

Marcot, B.G. and Penman, T.D. (2019), “Advances in Bayesian network modelling: Integration of modelling technologies”, *Environmental Modelling & Software*, Vol. 111, pp. 386–393, doi: 10.1016/j.envsoft.2018.09.016.

McNamara, G. and Bromiley, P. (1997), “Decision making in an organizational setting: cognitive and organizational influences on risk assessment in commercial lending”, *Academy of Management Journal*, Vol. 40 No. 5, pp. 1063–1088, doi: 10.2307/256927.

Mear, R. and Firth, M. (1988), “Risk Perceptions of Financial Analysts and the Use of Market and Accounting Data”, *Accounting and Business Research*, Routledge, Vol. 18 No. 72, pp. 335–340, doi: 10.1080/00014788.1988.9729380.

Mikes, A. (2009), “Risk management and calculative cultures”, *Management Accounting Research*, Vol. 20 No. 1, pp. 18–40, doi: 10.1016/j.mar.2008.10.005.

Mikes, A. and Kaplan, R.R.S. (2013), “Towards a Contingency Theory of Enterprise Risk Management”, *Harvard Business School*, pp. 1–47, doi: 10.2139/ssrn.2311293.

Milgram, S. (1963), “Behavioral Study of obedience”, *The Journal of Abnormal and Social Psychology*, American Psychological Association, US, Vol. 67, pp. 371–378, doi: 10.1037/h0040525.

Moeller, R.R. (2011), *COSO Enterprise Risk Management: Establishing Effective Governance, Risk, and Compliance Processes*, 2nd ed., Wiley, Hoboken, NJ.

Moreland, R.L. and Myaskovsky, L. (2000), “Exploring the performance benefits of group training: transactive memory or improved communication?”, *Organizational Behavior and Human Decision Processes*, Elsevier Science, Netherlands, Vol. 82 No. 1, pp. 117–133, doi: 10.1006/obhd.2000.2891.

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., *et al.* (2013), “Describing human decisions in agent-based models - ODD+D, an extension of the ODD protocol”, *Environmental Modelling and Software*, doi: 10.1016/j.envsoft.2013.06.003.

Nadkarni, S. and Narayanan, V.K. (2007), "Strategic schemas, strategic flexibility, and firm performance: the moderating role of industry clockspeed", *Strategic Management Journal*, Vol. 28 No. 3, pp. 243–270, doi: 10.1002/smj.576.

Nadkarni, S. and Shenoy, P.P. (2001), "A Bayesian network approach to making inferences in causal maps", *European Journal of Operational Research*, Vol. 128 No. 3, pp. 479–498, doi: 10.1016/S0377-2217(99)00368-9.

Nadkarni, S. and Shenoy, P.P. (2004), "A causal mapping approach to constructing Bayesian networks", *Decision Support Systems*, Vol. 38 No. 2, pp. 259–281, doi: 10.1016/S0167-9236(03)00095-2.

Neef, D. (2005), "Managing corporate risk through better knowledge management", edited by Firestone, J.*The Learning Organization*, Vol. 12 No. 2, pp. 112–124, doi: 10.1108/09696470510583502.

Neil, M., Fenton, N., Osman, M. and Lagnado, D. (2019), "Causality, the critical but often ignored component guiding us through a world of uncertainties in risk assessment", *Journal of Risk Research*, pp. 1–5, doi: 10.1080/13669877.2019.1604564.

Nielsen, S.H. and Parsons, S. (2007), "An application of formal argumentation: Fusing Bayesian networks in multi-agent systems", *Artificial Intelligence*, Vol. 171 No. 10–15, pp. 754–775, doi: 10.1016/j.artint.2007.04.005.

Nowak, G. and Thagard, P. (1992), "Newton, Descartes, and explanatory coherence", in Duschl, R.A. and Hamilton, R.J. (Eds.), *Philosophy of Science, Cognitive Psychology and Educational Theory and Practice*, State University of New York Press, Albany, New York, pp. 69–115.

O'Donnell, E. and Prather-Kinsey, J. (2010), "Nationality and differences in auditor risk assessment: A research note with experimental evidence", *Accounting, Organizations and Society*, Vol. 35 No. 5, pp. 558–564, doi: 10.1016/j.aos.2010.01.003.

Olson, D.L. and Wu, D. (2020), *Enterprise Risk Management Models*, Springer Berlin Heidelberg, Berlin, Heidelberg, doi: 10.1007/978-3-662-60608-7.

Ostrom, L. and Wilhelmsen, C. (2019), *Risk Assessment: Tools, Techniques, and Their Applications*, 1st ed., Wiley, doi: 10.1002/9781119483342.

Otway, H. and von Winterfeldt, D. (1992), "Expert Judgment in Risk Analysis and Management: Process, Context, and Pitfalls", *Risk Analysis*, Vol. 12 No. 1, pp. 83–93, doi: 10.1111/j.1539-6924.1992.tb01310.x.

Pan, Y., Siegel, S. and Wang, T.Y. (2017), "Corporate Risk Culture", *Journal of Financial and Quantitative Analysis*, Cambridge University Press, Vol. 52 No. 6, pp. 2327–2367, doi: 10.1017/S0022109017000771.

Paul, S. and Nazareth, D.L. (2010), "Input information complexity, perceived time pressure, and information processing in GSS-based work groups: an experimental investigation using a decision schema to alleviate information overload conditions", *Decision Support Systems*, Vol. 49 No. 1, pp. 31–40, doi: 10.1016/j.dss.2009.12.007.

Pearl, J. (2008), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Rev. 2., Kaufmann, San Francisco, CA.

Pearl, J. and Mackenzie, D. (2018), *The Book of Why: The New Science of Cause and Effect*, Basic Books, New York.

Pearl, J. and Russell, S. (2000), "Bayesian networks".

Porter, T.M. (1995), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, *Trust in Numbers*, Princeton University Press, doi: 10.1515/9781400821617.

Power, M. (2004), *The Risk Management of Everything: Rethinking the Politics of Uncertainty*, 1. publ., Demos, London.

Power, M. (2007), *Organized Uncertainty: Designing a World of Risk Management*, Oxford University Press, Oxford ; New York.

Power, M. (2016), *Riskwork. Essays on the Organizational Life of Risk Management*.

Power, M. (2020), "Risk Culture and Information Culture: Why an 'Appetite for Knowledge' Matters", in Ralph, D., Alexander, K. and Tuveson, M. (Eds.), *Beyond Bad Apples: Risk Culture in Business*, Cambridge University Press, Cambridge, pp. 42–72, doi: 10.1017/9781316996959.003.

Pritchard, C.L. (2015), *Risk Management: Concepts and Guidance*, Fifth edition., CRC Press, Taylor & Francis Group, Boca Raton.

Quail, R. (2011), "How to plan and run a risk management workshop", in Fraser, J. and Simkins, B.J. (Eds.), *Enterprise Risk Management*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 155–170, doi: 10.1002/9781118267080.ch10.

Railsback, S.F. and Grimm, V. (2011), *Agent-Based and Individual-Based Modeling: A Practical Introduction*, Princeton University Press, doi: 10.2307/j.ctt7sns7.

Rausand, M. (2011), *Risk Assessment: Theory, Methods, and Applications*, Wiley, Hoboken, N.J.

Rouse, W.B. and Morris, N.M. (1986), "On looking into the black box: prospects and limits in the search for mental models.", *Psychological Bulletin*, American Psychological Association, Vol. 100 No. 3, p. 349.

de Ruijter, A. and Guldenmund, F. (2016), "The bowtie method: A review", *Safety Science*, Vol. 88, pp. 211–218, doi: 10.1016/j.ssci.2016.03.001.

Schulz-Hardt, S., Brodbeck, F.C., Mojzisch, A., Kerschreiter, R. and Frey, D. (2006), "Group decision making in hidden profile situations: dissent as a facilitator for decision quality.", *Journal of Personality and Social Psychology*, Vol. 91 No. 6, pp. 1080–1093, doi: 10.1037/0022-3514.91.6.1080.

Secchi, D. (2015), "A case for agent-based models in organizational behavior and team research", *Team Performance Management: An International Journal*, Vol. 21 No. 1/2, pp. 37–50, doi: 10.1108/TPM-12-2014-0063.

Sheffield, J. (2004), "The design of GSS-enabled interventions: a Habermasian perspective", *Group Decision and Negotiation*, Vol. 13 No. 5, pp. 415–435, doi: 10.1023/B:GRUP.0000045750.48336.f7.

Shefrin, H. (2016), *Behavioral Risk Management*, Palgrave Macmillan US, New York, doi: 10.1057/9781137445629.

Simms, A. and Nichols, T. (2014), "Social Loafing: A Review of the Literature", *Journal of Management Policy & Practice*, Vol. 15 No. 1, pp. 58–67.

Simons, R. (1990), "The role of management control systems in creating competitive advantage: New perspectives", *Accounting, Organizations and Society*, Vol. 15 No. 1, pp. 127–143, doi: 10.1016/0361-3682(90)90018-P.

Sjöberg, L. (2000), "Factors in risk perception", *Risk Analysis*, Vol. 20 No. 1, pp. 1–12, doi: https://doi.org/10.1111/0272-4332.00001.

Son, J. and Rojas, E.M. (2011), "Evolution of Collaboration in Temporary Project Teams: An Agent-Based Modeling and Simulation Approach", *Journal of Construction Engineering and Management*, American Society of Civil Engineers, Vol. 137 No. 8, pp. 619–628, doi: 10.1061/(ASCE)CO.1943-7862.0000331.

Stasser, G. and Birchmeier, Z. (2003), "Group creativity and collective choice", in Paulus, P.B. and Nijstad, B.A. (Eds.), *Group Creativity: Innovation through Collaboration*, Oxford University Press, New York, Oxford, pp. 85–109.

Stasser, G. and Stewart, D. (1992), “Discovery of hidden profiles by decision-making groups: solving a problem versus making a judgment”, *Journal of Personality and Social Psychology*, American Psychological Association, US, Vol. 63 No. 3, pp. 426–434, doi: 10.1037/0022-3514.63.3.426.

Stasser, G. and Titus, W. (1985), “Pooling of unshared information in group decision making: biased information sampling during discussion”, *Journal of Personality and Social Psychology*, American Psychological Association, US, Vol. 48 No. 6, pp. 1467–1478, doi: 10.1037/0022-3514.48.6.1467.

Sun, L. and Erath, A. (2015), “A Bayesian network approach for population synthesis”, *TRANSPORTATION RESEARCH PART C-EMERGING TECHNOLOGIES*, Vol. 61, pp. 49–62, doi: 10.1016/j.trc.2015.10.010.

Sun, R. and Naveh, I. (2004), “Simulating organizational decision-making using a cognitively realistic agent model”, *JASSS*, Vol. 7 No. 3.

Sun, Z. and Müller, D. (2013), “A framework for modeling payments for ecosystem services with agent-based models, Bayesian belief networks and opinion dynamics models”, *Environmental Modelling and Software*, Vol. 45, pp. 15–28, doi: 10.1016/j.envsoft.2012.06.007.

Surowiecki, J. (2004), *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*, Doubleday & Co, New York, NY, US, pp. xxi, 296.

Thagard, P. (1989), “Explanatory coherence”, *Behavioral and Brain Sciences*, Vol. 12 No. 3, pp. 435–467, doi: 10.1017/S0140525X00057046.

Thagard, P. (2000), *Coherence in Thought and Action*, MIT Press, Cambridge, Mass.

Thagard, P. (2004), “Causal inference in legal decision making: Explanatory coherence vs. Bayesian networks”, *Applied Artificial Intelligence*, Vol. 18 No. 3–4, pp. 231–249.

Thagard, P. (2008), *Hot Thought: Mechanisms and Applications of Emotional Cognition*, Illustrated Edition., A Bradford Book, Cambridge, Mass.

Thagard, P. (2012), “Cognitive architectures”, in Frankish, K. and Ramsey, W.M. (Eds.), *The Cambridge Handbook of Cognitive Science*, Cambridge University Press, Cambridge, UK, pp. 50–70.

Thagard, P. and Verbeurgt, K. (1998), “Coherence as constraint satisfaction”, *Cognitive Science*, Vol. 22 No. 1, pp. 1–24.

Tommaso, P. (2017), : “Two sides of the same coin?”, *The Routledge Companion to Accounting and Risk*, Routledge.

Valacich, J.S., Sarker, S., Pratt, J. and Groomer, M. (2009), “Understanding risk-taking behavior of groups: A ‘decision analysis’ perspective”, *Decision Support Systems*, Vol. 46 No. 4, pp. 902–912, doi: 10.1016/j.dss.2008.12.003.

Vriezekolk, E., Etalle, S. and Wieringa, R. (2015), “Experimental Validation of a Risk Assessment Method”, in Fricker, S.A. and Schneider, K. (Eds.), *Requirements Engineering: Foundation for Software Quality*, Springer International Publishing, Cham, pp. 1–16, doi: 10.1007/978-3-319-16101-3_1.

Wall, F. (2016), “Agent-based modeling in managerial science: an illustrative survey and study”, *Review of Managerial Science*, Vol. 10 No. 1, pp. 135–193, doi: 10.1007/s11846-014-0139-3.

Wall, F. and Leitner, S. (2021), “Agent-based computational economics in management accounting research: opportunities and difficulties”, *Journal of Management Accounting Research*, Vol. 33 No. 3, pp. 189–212, doi: 10.2308/JMAR-19-073.

Weber, E.U. and Hsee, C. (1998), “Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk”, *Management Science*, INFORMS, Vol. 44 No. 9, pp. 1205–1217, doi: 10.1287/mnsc.44.9.1205.

Weber, P., Medina-Oliva, G., Simon, C. and Iung, B. (2012), “Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas”, *Engineering Applications of Artificial Intelligence*, Vol. 25 No. 4, pp. 671–682, doi: 10.1016/j.engappai.2010.06.002.

Wegner, D.M. (1987), “Transactive memory: a contemporary analysis of the group mind”, in Mullen, B. and Goethals, G.R. (Eds.), *Theories of Group Behavior*, Springer New York, New York, NY, pp. 185–208, doi: 10.1007/978-1-4612-4634-3_9.

Wolf, I., Schröder, T., Neumann, J. and de Haan, G. (2015), “Changing minds about electric cars: An empirically grounded agent-based modeling approach”, *Technological Forecasting and Social Change*, Vol. 94, pp. 269–285, doi: 10.1016/j.techfore.2014.10.010.

Wu, T., Huang, S., Blackhurst, J., Zhang, X. and Wang, S. (2013), “Supply Chain Risk Management: An Agent-Based Simulation to Study the Impact of Retail Stockouts”, *IEEE Transactions on Engineering Management*, presented at the IEEE Transactions on Engineering Management, Vol. 60 No. 4, pp. 676–686, doi: 10.1109/TEM.2012.2190986.

Yoe, C.E. (2019), *Principles of Risk Analysis: Decision Making under Uncertainty*, Second edition., Taylor and Francis, CRC Press, Boca Raton.

8 Appendix

Appendix 1 Glossary

Bayesian network	A probabilistic model that uses directed acyclic graphs (DAGs) to model the probabilities of events and their dependencies, based on the Bayesian theorem.
Benchmark	Here used as the best possible solution to the risk assessment task, i.e., the assessment an all-knowing individual would make.
Cognitive architecture	A framework that models the structure and functioning of cognitive processes in the human mind and explains how information is processed, stored, and retrieved.
CoMSES	A platform to share computer-based models and simulations.
Constraint satisfaction network	A model used to solve complex constraint problems by modelling variables and their dependencies and defining constraints that must be satisfied in order to find a solution.
ECHO	computational model proposed by Paul Thagard to explain the cognitive processes involved in analogical reasoning and problem solving.
Effectiveness	In the context of the simulation model: Ability of a risk workshop to provide correct risk assessments.
Facilitator	Here used as the person who organizes and moderates a risk workshop.
Information	Here used as a description of a fact that can be learned and shared.
Knowledge	Here used as the accumulation of facts and information that determine an individual's understanding.
Leader	Here used as the person who is responsible for the final decision on how severe a risk is.
Quantitative enthusiasm	Culture of passion and interest towards working with and analyzing numerical data and applying quantitative methods to gain insights and solve problems.

Quantitative skepticism	A culture with a critical attitude towards quantitative data, models and analyses and caution in drawing conclusions based solely on numerical information.
Risk assessment	The process of identifying, evaluating, and assessing potential risks in order to develop appropriate risk mitigation and control measures. Here used in the narrow sense of assessing risk severity.
Risk workshop	A structured discussion involving relevant stakeholders to identify, assess, and prioritize risks. Here used with a focus on risk assessment.
Participant	In the context of the simulation model: The individual stakeholders who participate in the risk workshop by sharing their knowledge and assessing the risk.
Transactive memory	A collective memory in which people use each other's knowledge together through knowing how information is distributed within a group.

Appendix 2 Additional results quantitative enthusiasm

Table 9 Results for simulation experiment 3 (group characteristics) after ten stable rounds.

Differences in the distribution of knowledge	Receivers consider hierarchical differences	Receivers have no transactive memory	Proportion of correct assessments			
			All risks (%)	High risks (%)	Low risks (%)	Avg. number of discussion rounds
+	+	+	65.7%	87.7%	43.3%	14.6
+	+	-	62.4%	81.3%	44.0%	14.9
+	-	+	63.7%	83.7%	44.3%	15.0
+	-	-	64.6%	82.8%	45.3%	14.5
-	+	+	75.1%	86.1%	64.7%	15.7
-	+	-	76.4%	87.2%	65.9%	15.7
-	-	+	75.7%	83.8%	67.4%	15.7
-	-	-	76.4%	87.2%	65.4%	15.7

Note: The table shows the percentage of risks that are correctly assessed after **ten stable discussion rounds** when the leader follows the majority vote and the average number of rounds before the decision is made, depending on group characteristics that might influence the risk workshop's effectiveness. A "+" indicates the presence of the corresponding deviation from an ideal discussion situation. Differences in the distribution of information are implemented by giving some participants a higher probability of receiving information during the initialization. If receivers consider hierarchical differences, they weigh input according to the sender's hierarchical position compared to their own. With no transactive memory, participants do not distinguish between senders who are experts on the information and those who are not.

Table 10 Results for simulation experiment 4 (interaction pattern) after ten stable rounds.

	Proportion of correct assessments			
Who talks next during the discussion?	All risks (%)	High risks (%)	Low risks (%)	Avg. number of discussion rounds
Random choice of participants	76.4%	87.2%	65.4%	15.7
Priority to concerned participants	69.9%	93.9%	47.6%	15.3
Priority to participants with dissenting opinions	74.2%	77.8%	70.4%	16.3
Priority to participants with higher hierarchical position	76.0%	86.4%	66.2%	15.6
Priority to participants close to group opinion	73.8%	93.2%	54.8%	14.9

Note: Percentage of risks that are correctly assessed after **ten stable discussion rounds** when the leader follows the majority vote and the average number of rounds before the decision is made, depending on the interaction pattern. Concerned participants are those who assess the risk as particularly high. Dissenting participants and those close to the group opinion are determined by measuring the distance between their risk assessment and the average risk assessment of the group.

Appendix 3 Sensitivity analysis on the number of participants

We have conducted a sensitivity analysis for the first study, focused on quantitative enthusiasm, regarding the number of participants for the workshop. Figure 20 depicts results of the sensitivity analysis that we conducted concerning the number of agents in a risk workshop.

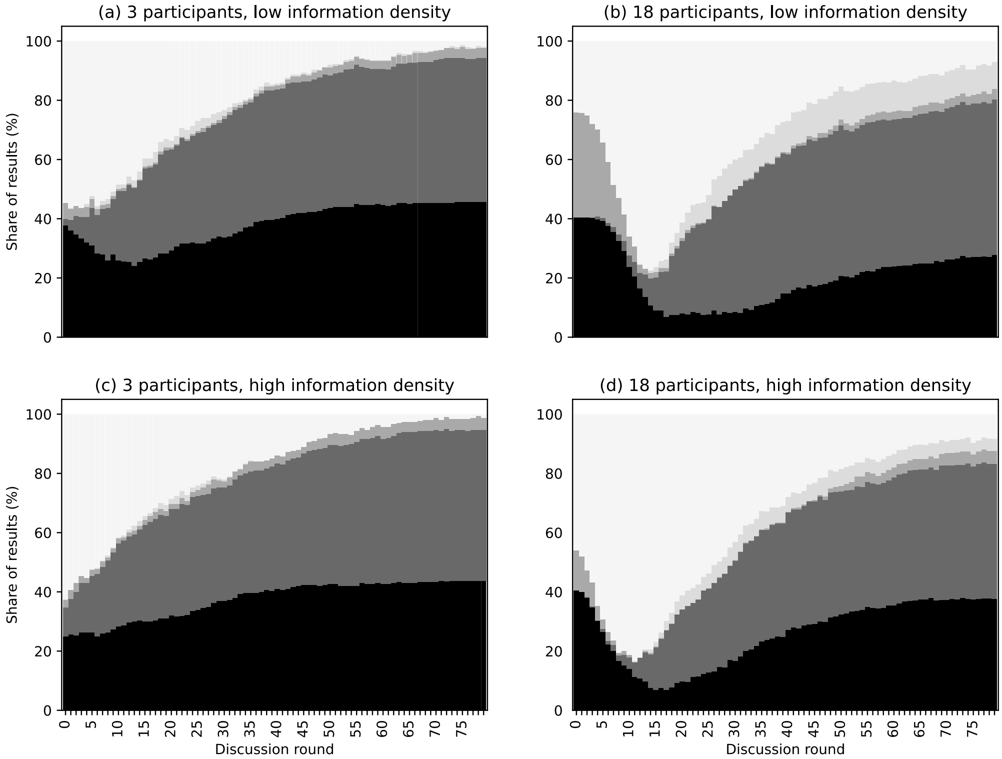


Figure 20 Comparison of model behavior using different numbers of participants and information density.

The experimental conditions and the analysis depicted in Figure 20 are analogous to the results for the first deviation reported in the study (limited information transfer). It shows the classification of the group assessment during the discussion.

Conducting a sensitivity analysis concerning the number of participants is not a straightforward exercise, as additional assumptions (1) about the density of information have to be made and (2) this also affects the knowledge of the agents about the risk structure. First, when the number of participants is changed, it has to be decided how much information is provided to each participant – is the same amount of information shared among a smaller or larger group, or is the amount of information per participant kept constant? Second, when information is only shared between 3 participants, they will have more knowledge about the risk structure from the beginning, so that the initial phase of learning about the structure is less pronounced.

In panel (b) and (c), the amount of information provided to the group as a whole is the same as in the paper. For 18 agents this results in a lower information density (panel (b)) while for three agents this implies a higher information density (panel (c)). These are qualitatively new settings compared to the study as for the lower information density (panel (b)) the learning task gets more challenging (because of less initially shared information due to low information density and less knowledge about the risk structure) while it is easier for the higher information density setting (panel (c)) (because of more initially shared information due to high information density and more knowledge about the risk structure). The graphs in panel (b) and (c) support this. However, that key qualitative patterns of the learning process remain still the same as in the paper, in the sense that participants already classify a substantial proportion of the low risk correctly and that the main proportion of the high risks is learned along the process. We also see for both settings, that in a later phase of the discussion the learning of high risks is at the cost of low risks, although this is less pronounced for the "easier" setting (panel (c)).

In panel (a), we depict results for a situation with three participants and a low information density, in panel (d) for 18 participants and a high information density. Key qualitative patterns of the process can again be identified. In both panels participants already classify a substantial proportion of the low risk correctly at the beginning and that the main share of the high risks is learned from the discussion. Again, for both settings the learning of high risks is at the cost of low risks in a later phase of the discussion, although less pronounced when the risk structure is better known from the beginning.

Overall, we see in these results support for our modelling choices concerning what problem the group is presented (e.g., in the sense that panel (c) would be too easy). More importantly, we observe still similar qualitative learning patterns across the variations we consider here with respect to key patterns. This collectively supports the assumption that the qualitative model behavior is relatively robust to the number of participants.

Appendix 4 ODD+D protocol quantitative enthusiasm

The model description follows the ODD+D protocol (Müller *et al.*, 2013), based on the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm *et al.*, 2006, 2020).[67]

ODD+D for the study on quantitative enthusiasm

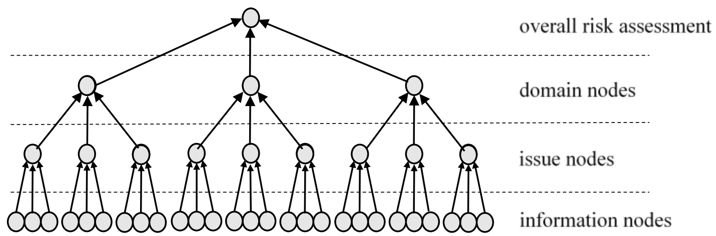
Outline		Guiding questions	ODD+D Model description
I) Overview	I.i Purpose	I.i.a What is the purpose of the study?	The study explores the drivers of the effectiveness of risk assessments in risk workshops, with reference to correctness and required time. Specifically, we model the limits to information transfer, incomplete discussions, group characteristics, and interaction patterns and investigate their effect on risk assessment in risk workshops.
		I.ii.b For whom is the model designed?	The model aims to guide facilitators of risk workshops in understanding the design choices and tradeoffs they face. The model also provides a way to use ABM for simulating complex individual cognition, which is valuable for scholars who simulate collaboration in organizations.
	I.ii Entities, state variables, and scales	I.ii.a What kinds of entities are in the model?	• Risk workshop participants, that discuss a risk in order to enable a correct risk assessment. • A leader, who is one of the participants in the risk workshop but makes the final decision. • A facilitator, who makes decisions regarding the proceedings of the workshop. • A risk, represented as a Bayesian network, that gets assessed in the risk workshop.

		I.ii.b By what attributes (i.e. state variables and parameters) are these entities characterized?	• Risk workshop participants (including leader):
			  o Initial information on risk
			    ▪ Participants are provided with some information about the risk during the initialization. Participants are considered experts regarding the information they are initially provided.
			  o Knowledge of the risk structure

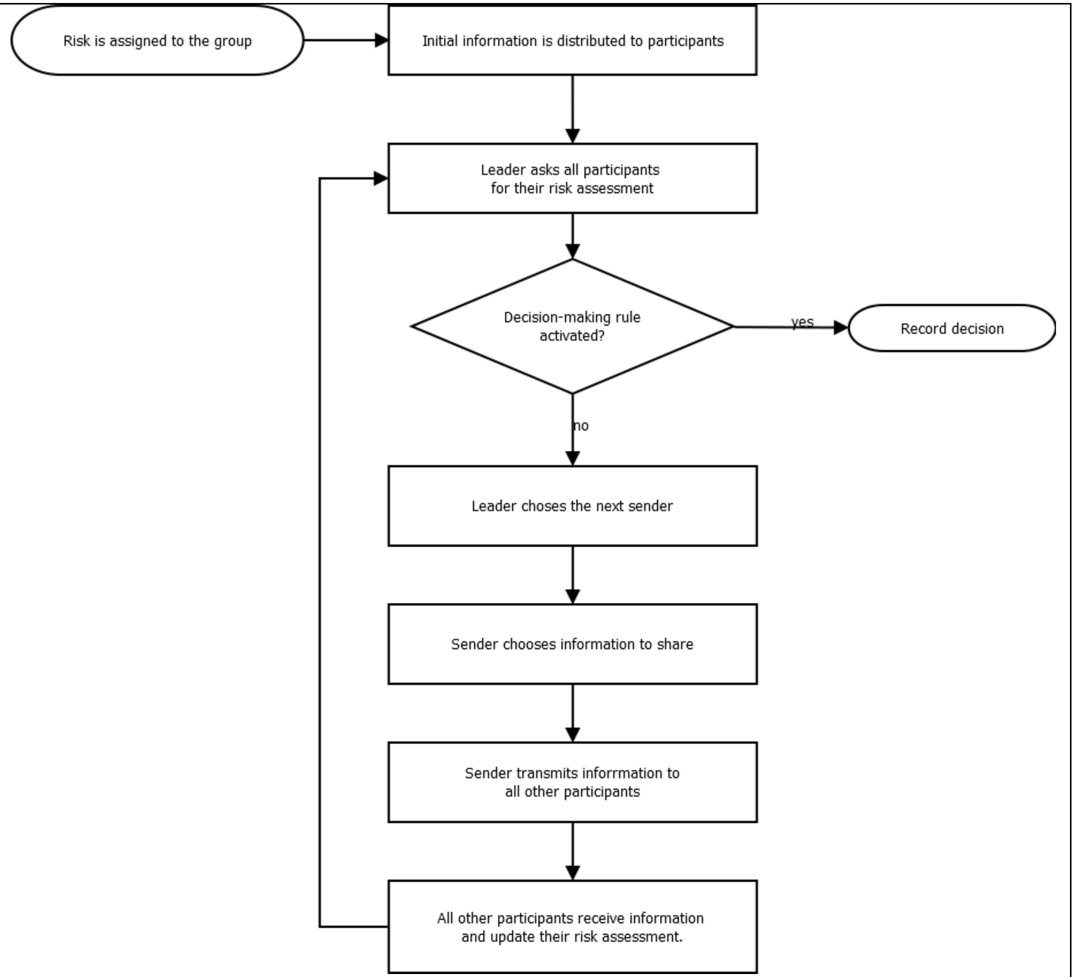
[67] This ODD+D protocol has been published along with the simulation code on CoMSES (Harten *et al.*, 2021).

    - ▪ Participants have a Bayesian network (Pearl, 2008) as their mental model of the risk. The Bayesian network of participants is initially limited to structures related to their expertise, (c.f. II.vi.c)
  - Level of hierarchy
    - ▪ Participants are randomly assigned one of three levels of hierarchy: low, medium, or high.
- Risk:
  - Risk structure
    - ▪ A risk is a Bayesian network with an overall risk assessment node and nodes representing 3 domains, 9 issues (3 per domain) and 27 information (3 per issue).
    - ▪ Each node has three states (low, medium, high) with an associated likelihood for each state (see e.g. Kabir *et al.*, 2015 for a similarly structured risk network).
  - information about the risk
    - ▪ Each of the 27 information nodes has a true state (low, medium, or high). The true states of the information nodes are used to calculate the benchmark risk assessment.

See the following two sections after the ODD+D protocol for an explanation of how the Bayesian network is calibrated and for examples of how the overall risk assessment is derived from the states of the information nodes.



		I.ii.c What are the exogenous factors / drivers of the model?	Several experiments are conducted with the model. Depending on the experiment, the following attributes are systematically varied: • Initial distribution of information among the risk workshop participants o Either all participants get the same amount of information initially, or information is assigned unequally initially. • Consideration of hierarchy by the participants • Presence of transactive memory within the group • Decision rules used by the leader to make decisions • Interaction pattern within the group
		I.ii.d If applicable, how is space included in the model?	Space is not included in the model.
		I.ii.e What are the temporal and spatial resolutions and extents of the model?	One time step is one discussion round, meaning that a participant is chosen to speak, the speaker provides an information, and all other participants process the information provided to them. See I.iii.a for the steps that happen during one discussion round. The discussion is simulated for 140 rounds of discussion.
	I.iii Process overview and scheduling	I.iii.a What entity does what, and in what order?	The participants repeatedly share information with each other about the risk. The facilitator chooses one participant as a sender. All other participants become receivers. The discussion continues until the decision-making rule is activated.

			
II) Design		II.i.a Which general concepts, theories or hypotheses are underlying the	The model has been developed to investigate the impact of choices made during the facilitation of a risk workshop. While some results are specific to decision-making regarding risk assessment, the

Concepts		model's design at the system level or at the level(s) of the submodel(s) (apart from the decision model)? What is the link to complexity and the purpose of the model?	general concept can be transferred to other settings where a group shares distributed knowledge in a discussion to make a decision. The experiments concern deviations from an ideal speech situation (Habermas, 1982) which serves as a baseline for the effectiveness of the decision-making process. See Chapter 4.2.2 for a discussion of the deviations included in the model. The interaction of the participants is built upon the idea of participants forming a transactive memory system (Wegner, 1987): All information is available to the group, but the group needs to correctly make use of the available information. For this, it is essential to assess the reliability of information provided by other participants, e.g., based on knowledge about their expertise.
	II.i Theoretical and Empirical Background	II.i.b On what assumptions is/are the agents' decision model(s) based?	The participants are either following simple mathematical formulas for their decisions, or make random choices, where the probability of each option is defined by model parameters.
		II.i.c Why is a/are certain decision model(s) chosen?	The participants face decisions with a limited set of options (e.g., which participant should be the next sender). As there is no clear correct decision, we assume that each decision should be possible within the model, however the probability of each choice should be influenced by reasonable heuristics (e.g., when the group is aware of hierarchy, the facilitator should be more likely to choose agents with a high position in the hierarchy as the next senders).
		II.i.d If the model / a submodel (e.g., the decision model) is based on empirical data, where does the data come from?	The model is not based on empirical data.

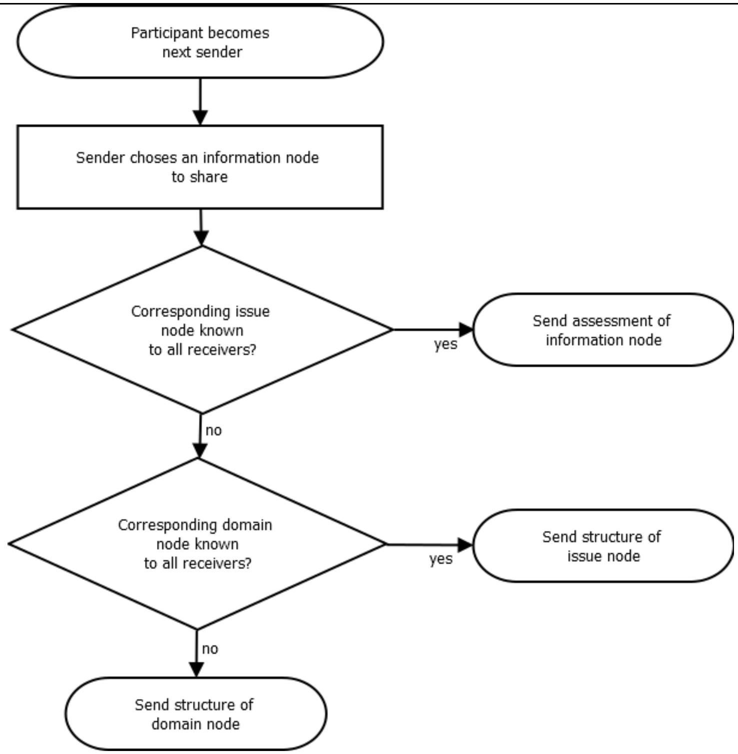
		II.i.e At which level of aggregation were the data available?	Not applicable.
	II.ii Individual Decision Making	II.ii.a What are the subjects and objects of decision-making? On which level of aggregation is decision-making modeled? Are multiple levels of decision making included?	The facilitator makes one decision: • Each round, one participant is chosen by the facilitator to be the next sender. The leader makes two decisions: • After each round, the leader decides if the conditions for a decision rule are met and therefore a decision can be made. • The leader also makes the final decision on how to assess the risk. Participants make two types of decisions: • What information to talk about when they are chosen to be senders for a discussion round. • How to weigh the input they receive from the sender if they are chosen to be receivers for a discussion round.
		II.ii.b What is the basic rationality behind agents' decision-making in the model? Do agents pursue an explicit objective or have other success criteria?	Participants make their best effort to gain a correct understanding of the risk, in order to reach an accurate risk assessment. The objective of the leader is to reach an accurate risk assessment in the shortest possible time.
		II.ii.c How do agents make their decisions?	Participants make their decisions by random choice; however, the probability of each decision might not be equal. E.g., participants are more likely to talk about information they are experts on.

		II.ii.d Do the agents adapt their behavior to changing endogenous and exogenous state variables? And if yes, how?	They do not.
		II.ii.e Do social norms or cultural values play a role in the decision-making process?	The selection of the next sender is influenced by social norms (e.g., prioritize participants based on hierarchy) in some experimental settings.
		II.ii.f Do spatial aspects play a role in the decision process?	No.
		II.ii.g Do temporal aspects play a role in the decision process?	The leader might decide to end the discussion and make a decision if the group has not progressed for some time.
		II.ii.h To which extent and how is uncertainty included in the agents' decision rules?	The participants can take uncertainty regarding the correct risk assessment into account, e.g. when the facilitator chooses the next sender.
	II.iii Learning	II.iii.a Is individual learning included in the decision process? How do individuals change their decision	The agents understanding of the risk under assessment is modelled as Bayesian network. Participants update their understanding of the risk, both concerning the risk structure and information about the risk, based on the input they receive during the discussion from other participants. Their understanding of the risk determines their individual risk assessment. They do not learn beyond the discussion of an individual risk, i.e. there is no interaction over several runs of the simulation.

		rules over time as consequence of their experience?	
		II.iii.b Is collective learning implemented in the model?	By exchanging their individual knowledge, the participants' understandings of the risk move towards a common understanding of the risk.
	II.iv Individual Sensing	II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? Is the sensing process erroneous?	None.
		II.iv.b What state variables of which other individuals can an individual perceive? Is the sensing process erroneous?	Depending on experimental conditions, participants know other participants' relative position in the hierarchy and other participants' expertise regarding topics. This knowledge, if available, is not erroneous. The facilitator and the leader know the overall risk assessment of all participants and can chose the next sender and decide when to end the discussion based on this knowledge (see also the flowchart in I.iii.a). Knowledge and information about the risk is explicitly exchanged in a simulated discussion.
		II.iv.c What is the spatial scale of sensing?	Not applicable.

		II.iv.d Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables?	Participants are assumed to know variables like the expertise of other participants (if the corresponding experimental condition is present). Participants learn about other participants' understanding of the risk via the simulated discussion.
		II.iv.e Are costs for cognition and costs for gathering information included in the model?	The cost of cognition and information gathering is accounted for by making a decision as soon as specific decision criteria are met, instead of continuing the discussion potentially infinitely. However, the participants make no conscious decision on whether to invest cognitive resources: whenever they get new input, they update their knowledge about the risk and their risk assessment.
	II.v Individual Prediction	II.v.a Which data uses the agent to predict future conditions?	Participants do not predict future conditions.
		II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?	Not applicable.
		II.v.c Might agents be erroneous in the prediction process, and how is it implemented?	Not applicable.

	II.vi Interaction	II.vi.a Are interactions among agents and entities assumed as direct or indirect?	The participants directly interact with each other by exchanging information about the risk during the discussion.
		II.vi.b On what do the interactions depend?	In each discussion round, one participant is chosen to talk to all other participants.
		II.vi.c If the interactions involve communication, how are such communications represented?	Participants communicate by exchanging information about their individual knowledge about the risk. Usually, they will share the activation of each state of the information node they have decided to share. However, initially participants do not know the full structure of the risk Bayesian network. They are only aware of information nodes provided to them, direct siblings of these information nodes, and all issue and domain nodes that are (grand)parents of these information nodes. Therefore, if not all receivers are aware of the existence of an information node the sender wants to talk about, the sender will use the discussion round to communicate information about the risk structure instead:

			
		II.vi.d If a coordination network exists, how does it affect the agent behaviour? Is the structure of the network imposed or emergent?	The setting for the simulation is one discussion by all participants. Therefore, each participant can send information to all other participants, if chosen to be the sender.

	II.vii Collectives	II.vii.a Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Are these aggregations imposed by the modeller or do they emerge during the simulation?	The individual participants form a group. The risk assessments of the individuals can be aggregated to a group opinion, like a consensus, a majority vote, or an average vote (used to determine if the group opinion is moving over time). The leader makes the decision to end the discussion based on such aggregates.
		II.vii.b How are collectives represented?	The collective has no agency by itself and is only a conceptual component of the model.
	II.viii Heterogeneity	II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?	The group participants differ in their initial knowledge and expertise. They might differ regarding their position in a hierarchy if the corresponding experimental condition is present. Notably, the participants initially have heterogeneous mental models, as the risk structure of their mental models depends on the information provided to them.
		II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?	They are not heterogeneous in their decision-making.
	II.ix Stochasticity	II.ix.a What processes (including initialization) are modeled by assuming they	• For each run, a randomly generated risk is chosen for the group to assess. • The decision by the facilitator of who is the next sender in a discussion round is (partly) random. • The decision by the sender what to share is (partly) randomly.

		are random or partly random?	
	II.x Observation	II.x.a What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected?	At the beginning of each run, the independent variables are recorded. After each discussion round, the activation of all states of the 'overall assessment' node is recorded for each participant. Furthermore, it is checked and recorded if a decision rule (see Chapter 4.3.2) has been activated during the most recent discussion round, and what decision by the leader it would result in.
		II.x.b What key results, outputs or characteristics of the model are emerging from the individuals? (Emergence)	We can observe how limitations to the ideal speech situation impact the effectiveness of risk workshops, measured as the share of high and low risks that get assessed correctly, as well as the time it takes to make a decision.
III) Details	II.i Implementation Details	III.i.a How has the model been implemented?	The simulation is written in Python. The *pgmpy* library is used to create risks by sampling from the reference Bayesian network. The Bayesian networks are calculated in R using the *bnlearn* library.
		III.i.b Is the model accessible and if so where?	The model will be published on the OpenABM platform.
	III.ii Initialization	III.ii.a What is the initial state of the model world, i.e., at time t=0 of a simulation run?	The benchmark assessment The simulation requires a benchmark assessment that assessments reached by the workshop can be compared against. The benchmark assessment is determined by calculating a Bayesian network with full information about the true risk. This is the assessment that a group would reach given an ideal speech situation (the benchmark process). <u>Initial information distribution</u>

			In the beginning, information on the risk is provided to the participants. After this initial seeding of (true) information, participants only get new input from other participants. All participants are able to assess the overall risk based on the limited knowledge they are provided initially. Each information about the risk is available to at least one participant so that, in an ideal speech situation, a correct risk assessment is achievable. *initial_distr_info* bits of information are distributed among the participants, either equally or unequally (see III.ii.b). <u>Knowledge about each other's' hierarchical position</u> Each participant is assigned one of three hierarchical statuses (low, medium, high). If the corresponding experimental condition is present, participants are aware of the hierarchical status of the other participants and consider it when including sender input into their mental model. <u>Knowledge about each other's expertise</u> The information provided to the participants initially constitutes their expertise regarding the risk. Depending on the experimental condition, participants might be informed about each other's expertise, i.e., who is an expert on which information.
		III.ii.b Is initialization always the same, or is it allowed to vary among simulations?	<u>Risk task provided to the participants</u> A new risk is generated for each simulation run and information is assigned to the participants in a randomized process. <u>Initial information distribution</u>

			Depending on the experimental condition, information is provided equally or unequally to the participants. If the information is provided equally, each participant has the same probability of receiving any bit of information. If the information is provided unequally, the probability of participant i (out of n participants) to be provided with information j is: $$P(i,j,n) = \frac{2^i}{2^n - 1}$$ <u>Knowledge about each other's hierarchical position</u> Depending on the experimental condition, participants are provided with information about each other's position in the hierarchy, allowing the participants to consider it during the inclusion of new information. <u>Knowledge about each other's expertise</u> Depending on the experimental condition, participants are provided with information about each other's expertise, allowing the participants to consider it during the inclusion of new information.
		III.ii.c Are the initial values chosen arbitrarily or based on data?	The initial values are chosen arbitrarily, within the constraints set by the model parameter.
	III.iii Input Data	III.iii.a Does the model use input from external sources such as data files or other models to represent processes that change over time?	The model does not use external sources for input.

	III.iv Submodels	III.iv.a What, in detail, are the submodels that represent the processes listed in 'Process overview and scheduling'?	<u>Sender selection</u> The sender is chosen by the facilitator in a random draw that depends on the experimental condition. Only agents who have any expertise in any node are eligible to become senders, for technical reasons. See chapter 3.3.4 for the different interaction pattern investigated. Random: In the baseline mode, the next participant to speak is chosen at random, with an equal probability for each participant. Priority given to concern: The probability of each participant to be chosen is weighted by the value a participant assigns to the 'high' state of the overall risk assessment node. Priority given to dissent: The probability of each participant to be chosen is weighted by the distance of their individual risk assessment from the average group risk assessment. This distance is calculated as the sum of the absolute differences between the low-, medium-, and high states of the overall risk assessments by the individual participants and the average group risk assessment. Priority given to hierarchy: Participants are more likely to be the sender if they are assigned a higher hierarchical position. The probability of each participant to be chosen is weighted by *weight_h_low*, *weight_h_medium*, or *weight_h_high*. Priority given to homogeneity: The probability to be the next sender is higher if the participant's risk assessment is close to the average group risk assessment. The distance is calculated in the same way as when priority is given to dissent, however the probability of each participant to be chosen is weighted by the reciprocal of the distance to the group assessment. <u>Selection of sender output</u> The senders chose one of the information nodes available to them. The chance of choosing a specific node is weighted by the expertise the sender has regarding the information node. The expertise

			is *weight_e_expert* for nodes initially assigned to the sender, and *weight_e_nonexpert* for all other nodes. The sender will provide the detailed state of the information node to all other participants, who are the receivers in the given round. Furthermore, participants prioritize talking about nodes that have not been talked about often before. The previously determined weight for each node ($node_i$) based on the participant's expertise regarding the node is multiplied by a weight reflecting how much it has been talked about by the group already, relative to the node that was talked about most often ($node_{max}$): $weight(node_i) = weightExp(node_i) * 2^{discussionCounter(node_{max}) - discussionCounter(node_i)}$ If the sender wants to transmit an information node that is not part of the network of all receivers, the sender will instead send information on the first parent node that is not available to all receivers (see II.vi.c). <u>Inclusion of sender input</u> All receivers update their believes about the node the sender talks about by assigning the input a weight relative to their own, prior believe. Depending on the experimental condition, the weight assigned to the input reflects the following aspects: • difference in expertise regarding the specific node between sender and receiver • difference in hierarchy between sender and receiver – E.g., the receiver will weigh the input higher if the sender is an expert and the receiver is not, or if the sender has a higher position in the hierarchy compared to the receiver.

			For example, if a non-expert assesses the likelihood of an information node to be high as 0% and received the information of an expert that assesses the likelihood to be 23%, the updated belief of the non-expert will be – $23\% * (\text{weight\_e\_expert} - \text{weight\_e\_nonexpert}) - 0\% * (1 - (\text{weight\_e\_expert} - \text{weight\_e\_nonexpert}))$ The model also allows to weight the input based on the trust the receiver has in the sender. For this study, however, the trust is not varied and identical between all participants.
		III.iv.b What are the model parameters, their dimensions and reference values?	See A.2
		III.iv.c How were submodels designed or chosen, and how were they parameterized and then tested?	The submodels where designed to reflect a prototypical implementation of the processes that happen during a risk workshop. The experimental conditions were derived as deviations from an ideal speech situation and implemented as simple as possible. All decisions made by agents are either random or influenced by simple heuristics.

Parameters of the model

Parameter	Explanation	Default value(s)
nr_agents	Number of participants	9
rounds_max	Maximum number of simulated rounds	140
theshold_change continue	Maximum deviation in the group risk assessment for the discussion to be considered stagnant.	0.01
Risk model		
BN_nr_children	Number of children each parent-node has in the reference risk network	3
BN_factor_lmr	Factor for how much more likely the true state of an information is to be low rather than medium or medium rather than high	10
Participants		
weight_h_low	Weight used for weighting the input from senders if hierarchy is to be considered	0.25
weight_h_medium		0.50
weight_h_high		0.75
weight_e_expert	Weight used for weighting the input from sender if transactive memory is present	1.0
weight_e_nonexpert		0.1
initial_trust	The model allows to vary the trust participants have in each other. (Note: for the experiments included in the study, trust is kept constant at the default value)	1

initial_distr_info	Number of information to be distributed initially	45
Experimental conditions		
information_distribution	Determines whether all agents have the same probability to receive an information in the initial distribution	equal / unequal
weight_hierarchy	Participants consider the difference in hierarchy when weighting sender input	yes/no
weight_trans_mem	Participants consider the difference in expertise when weighting sender input	yes/no
mode_limited_trans	Participants consider their prior belief when integrating new information	yes/no
decision_rule	Decision-making rule used by the leader	See chapter 3.3.2
Interaction_pattern	Interaction pattern used in the group	See chapter 3.3.4

Calibration of the Bayesian network

The Bayesian network that is used to model the risk needs to be calibrated to provide a plausible risk assessment from the information that constitute the risk assessment task.

The overall risk assessment is derived from the state of the overall risk assessment node (see Figure 21) based on the highest associated probability. We have consciously chosen a calibration for the Bayesian network that will always result in a "high" or "low" overall risk assessment. In the Bayesian network used in this study, from aggregation level to aggregation level the likelihood of "medium" states decreased strongly making it de facto a binary overall risk. Overall, this makes it possible to clearly identify the assessments made as a result of the risk workshop as correct or incorrect, streamlining the communication and discussion of the simulation results.

When Bayesian networks are used for real-world applications, the calibration of the network can be derived using machine learning algorithms on a real-world dataset. Lacking such real-world data, we make plausible assumptions on how the network should aggregate information provided at the information nodes. Technically, the aggregation nodes (the topic nodes, the domain nodes, and the overall assessment node) can be thought of as a lookup table that provides a probability value for the 'low', 'medium,' and 'high' state based on the probabilities of states that feed into the aggregation node.

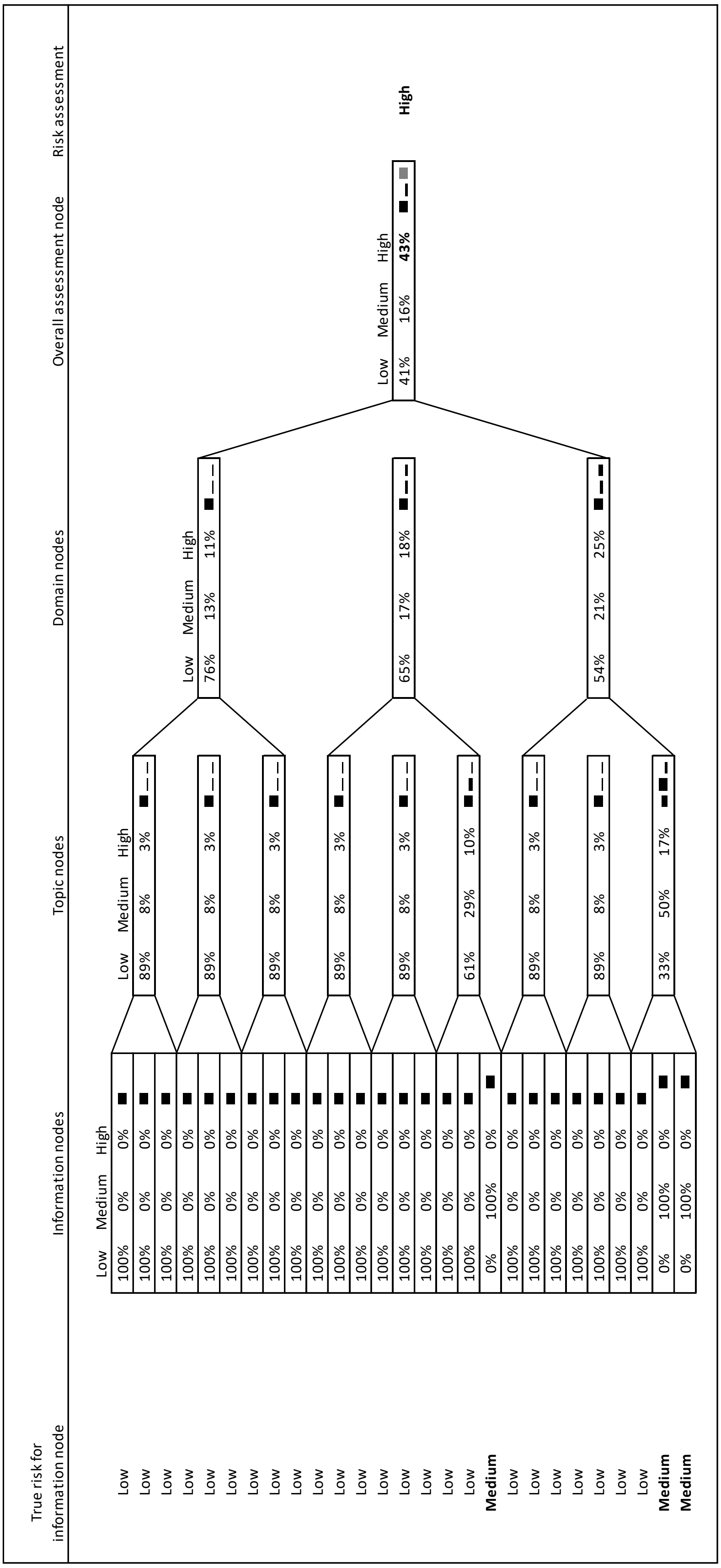


Figure 21 Example of how the BN calculates the overall risk assessment as a benchmark.

Figure 21 provides an overview of how a benchmark assessment is calculated. It is also an example of how, even with some ‘medium’ information provided, the ‘low’ and ‘high’ states of the overall assessment nodes have higher probabilities than the ‘medium’ state. While for one topic node (the last one at the bottom of Figure 21), the ‘medium’ state has the highest probability, at each further level, more information is aggregated, and the ‘low’ and/or ‘high’ states become more dominant.

The aggregation of the nine values that feed into each aggregation node (three for each connected node, with one likelihood for the low state, one for the medium, and one for the high) is implemented in a two-step process:

First, the input is aggregated into a “risk score” in the range 0 to 1. Here is an example from nodes in Figure A.3.1: This example is the case in which all three nodes to be aggregated carry a “low risk” information (e.g., the aggregation of the first three information nodes at the top of Figure A.3.1). A large probability in the ‘low’ states will result in a low risk score.

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 0.01 & 0.09 & 0.90 \end{bmatrix} * \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = [0.01]$$

Second, this aggregate risk score (0.01) is now passed to a lookup table that translates each risk score into probability values for all three states. Figure A.3.2 shows how the aggregate risk score translates into state probabilities. The values were chosen to reach plausible aggregates for some plausible inputs. For example, an input of ‘medium’ on all input nodes will provide the highest value for ‘medium’ at the aggregate node. However, it was important to calibrate the network so that even a single high probability for a ‘high’ state translates into a high probability of the ‘high’ state of the aggregate node, as we assume that most information nodes are always in the ‘low’ state, and the task of the risk workshop is to correctly handle a low number of ‘high risk’ or ‘medium risk’ information.

The configuration of the aggregate node that is derived this way for the topic nodes is equally applied for all aggregate nodes.

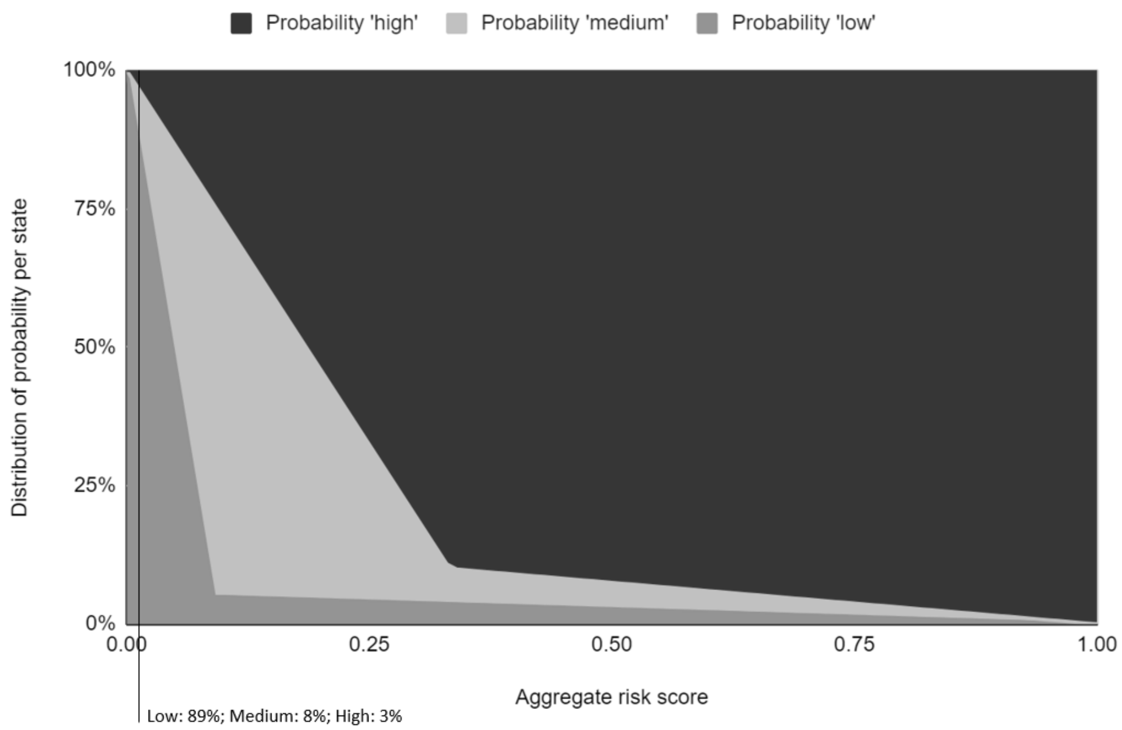


Figure 22 Graph on the translation of the aggregate risk score into state probabilities

Information processing in the Bayesian network

The mental model of the participants, implemented as a Bayesian network, allows to derive an overall risk assessment from a set of beliefs about information relevant to the risk. If all information is provided without any noise (as in the end of a discussion with ideal speech situation), the assessment of a participant is identical to the benchmark assessment. The participant is right for the right reasons (Figure 23).

The participant can also reach a correct assessment with much more limited information (Figure 24). However, in this case the assessment is correct only due to uncertainty about information: The ‘high’ information is not available to the participant – the participant is right for the wrong reason.

As the participant gains access to more information (again, assuming a unhindered discussion as in the ideal speech situation), the overall assessment of the participant switches to ‘low’ (Figure 25). By learning more (correct) information, the assessment becomes wrong.

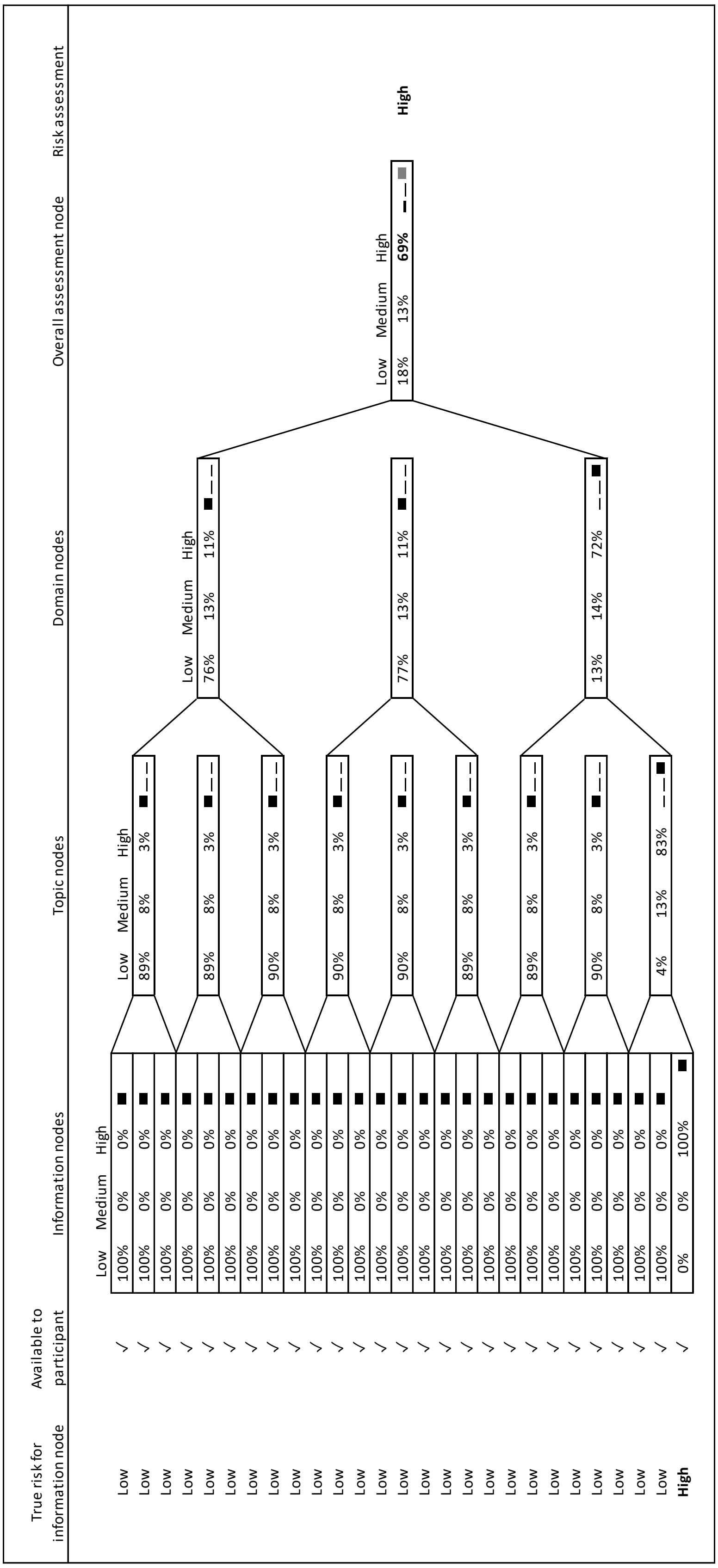


Figure 23 Correct assessment with all information

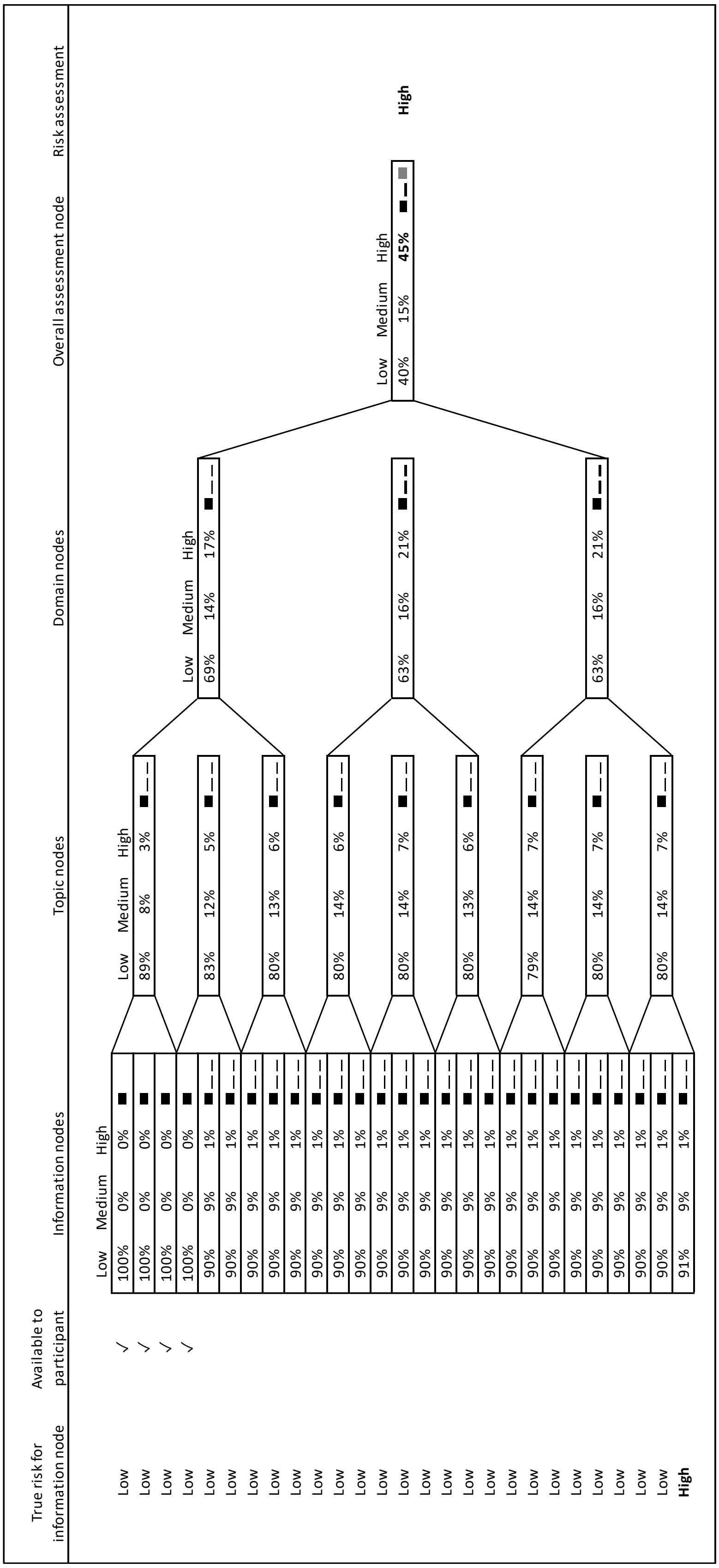


Figure 24 Correct assessment with few information

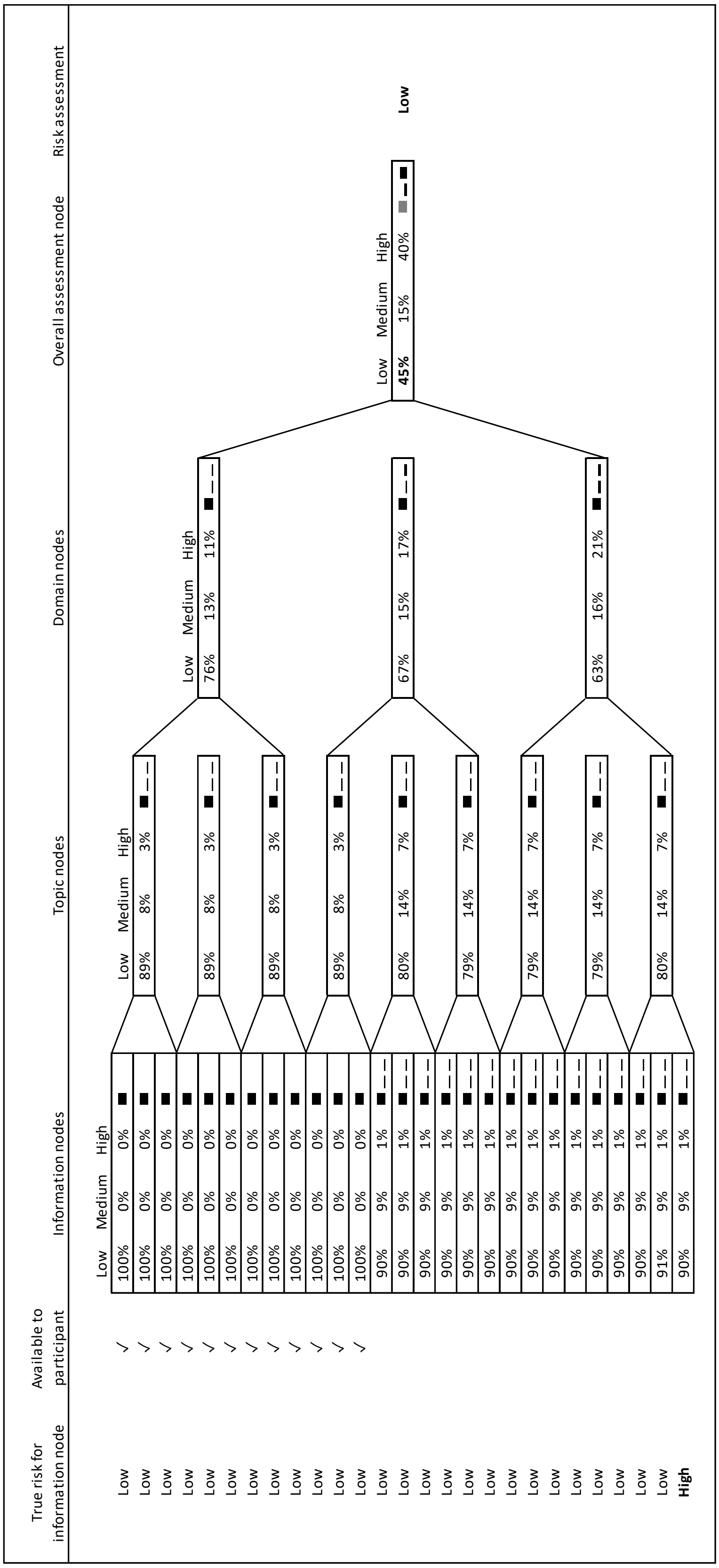


Figure 25 Wrong assessment with more information

Appendix 5 ODD+D protocol quantitative skepticism

The model description follows the ODD+D protocol (Müller *et al.*, 2013), based on the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm *et al.*, 2006, 2020). [68]

ODD+D for the study on quantitative skepticism

Outline		Guiding questions	ODD+D Model description
IV) Overview	I.i Purpose	I.i.a What is the purpose of the study?	The study builds upon Harten *et al.* (2022), which is a theoretical exploration of the drivers of the effectiveness of risk assessments in risk workshops regarding the correctness and required time. Harten *et al.* (2022) model agent's cognition using Bayesian networks. We argue that their modeling choice corresponds well to an organizational environment where a calculative culture of quantitative enthusiasm is prevalent. The purpose of our study is to investigate the impact of calculative cultures on risk workshops. More specifically, we explore drivers of the effectiveness of risk assessments in risk workshops dominated by 'quantitative skepticism' and contrast our findings with previous research that assumed the dominance of 'quantitative enthusiasm.' To this end, we adapt their model to a setting that better corresponds to an environment where quantitative skepticism is prevalent. We, therefore, chose constraint satisfaction networks. Specifically, we use ECHO networks (Thagard, 1989) with

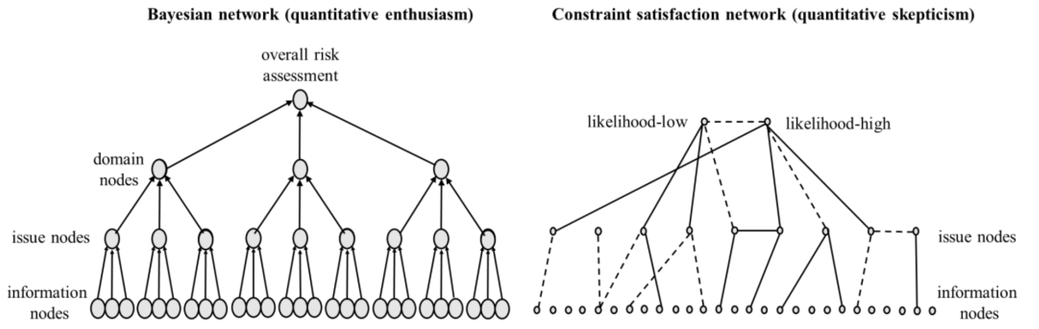
[68] This ODD+D protocol has been published along with the simulation code on CoMSES (see Bellora-Bienengräber *et al.*, 2022).

			a similar structure to the Bayesian networks used by Harten *et al.* (2022) to model agents' cognitive processes. In order to be able to compare results with Harten *et al.* (2022), we perform the same experiments as described in the original paper and the corresponding ODD (available at Harten *et al.*, 2021). Specifically, we model the limits to information transfer, incomplete discussions, group characteristics, and interaction patterns and investigate their effect on risk assessment in risk workshops.
		I.ii.b For whom is the model designed?	The model ideally guides facilitators of risk workshops in understanding the design choices and trade-offs they face, especially the importance of the prevalent calculative culture for accounting and design-related decisions. The model also provides a blueprint to use ABM to simulate organization discussion processes.
	I.ii Entities, state variables, and scales	I.ii.a What kinds of entities are in the model?	This ABM has nine agents, all of whom are participants in a risk workshop. In each simulation run, one risk task is generated in the form of a constraint satisfaction network. The participants perform a discussion, during which they exchange information about the risk. During the discussion, one agent (chosen at random) takes a special role as a leader, in addition to the role as a participant. Additionally, a facilitator's role is needed during the discussion process. However, - as we detail later - the role of the facilitator only requires information

			available to all participants or any external observer; the role of the facilitator is not implemented as an agent entity itself but only as a function. Overall, the following entities are present in the model: Risk workshop participants who discuss a risk in order to reach a correct risk assessment; a leader who is one of the participants in the risk workshop but is responsible for making the final decision; a facilitator, who makes decisions regarding the proceeding of the workshop (Should the discussion continue, or a decision be made? Who is the next participant to speak?); a risk, modeled as a constraint satisfaction network that is assessed in the risk workshop.
		I.ii.b By what attributes (i.e. state variables and parameters) are these entities characterized?	Risk workshop participants (including the leader) are characterized by three attributes: Initial information concerning the risk: Participants are provided with some information about the risk during the initialization. The pool of information that is assigned to the participants corresponds to the information nodes in the figure below.

			If participants are provided information, they integrate the corresponding information node into their constraint satisfaction network. Participants are considered experts regarding the information they are initially provided. Knowledge of the risk structure: Participants have a constraint satisfaction network as their mental model of the risk. The constraint satisfaction network (Thagard, 1989) of the participants is initially limited, as some information nodes and their relationship to other nodes are missing. After participants receive new information, they update their individual constraint satisfaction network by adding new nodes and their relationships with the nodes already in their network (either explanatory or contradictory). Level of hierarchy: Participants are randomly assigned to one of three levels of hierarchy: low, medium, or high. There are always three participants for each level of the hierarchy. The leader is randomly chosen from the participants with a high level of hierarchy. The risk under assessment is characterized by the relationships between the nodes of the full constraint satisfaction network. A risk is a constraint satisfaction network with two nodes representing the risk assessments “low likelihood risk” and “high likelihood risks” and nodes representing nine issues and 27 information. Each node has an activation in the range between -1 and 1, representing the strength of belief in the corresponding node. The activation changes depending on

the relationship of the nodes with each other. This (symmetric) relationship is either explanatory or contradictory (or there is no direct relationship at all). For example, a piece of specific information ("a new competitor enters the market with a similar product") might have an explanatory relationship with one issue ("the market we operate in is attractive") node and a contradictory relationship with another issue node ("we have a technological lead in the market").

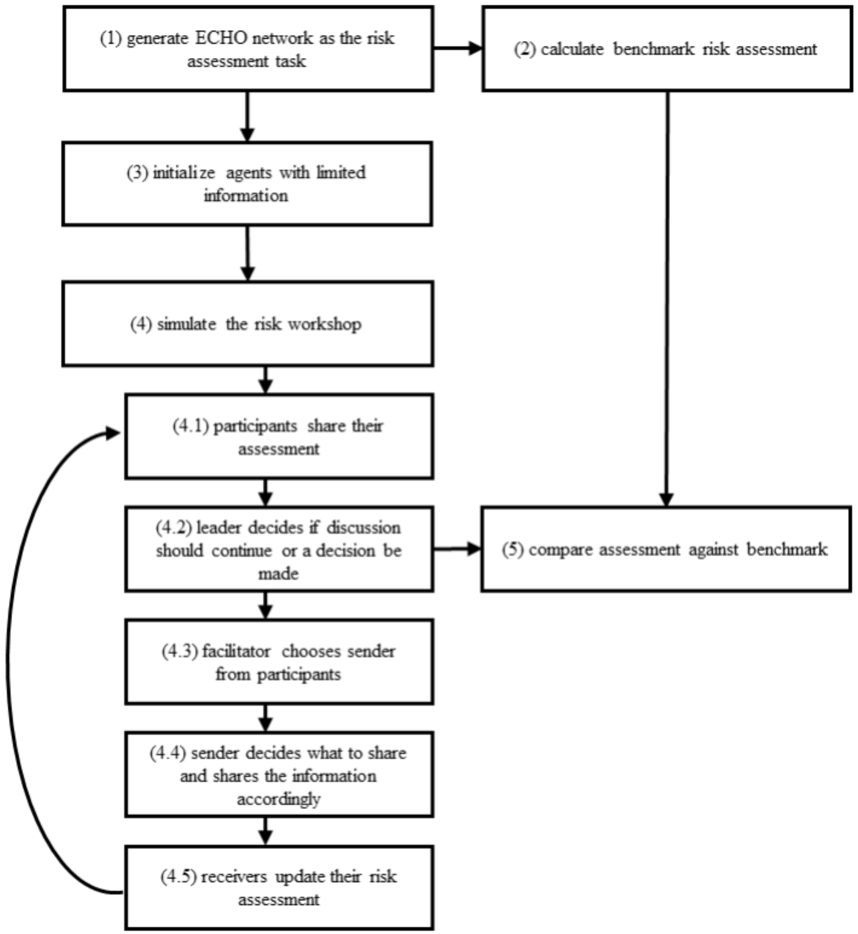


The structure of the Bayesian network used by Harten *et al.* (2022) (left) and the stylized ECHO network used to model quantitative skepticism (right).

The benchmark risk assessment is derived from the activation of the two risk assessment nodes depicted at the top when all information nodes are provided instantly (e.g., the risk assessment of an agent with access to the whole constraint satisfaction network from the beginning).

		I.ii.c What are the exogenous factors / drivers of the model?	Several simulation experiments are conducted with the model. Depending on the experiment, the following attributes are systematically varied: • Initial distribution of information among the risk workshop participants: Either all participants get the same amount of information, or information is assigned unequally. Specifically, each information from the pool of information that is initially assigned to the group is distributed one after another. For each information, each participant has a specific probability of receiving this information. If information is distributed equally, each agent's probability of receiving a piece of specific information is 11.1% (i.e., 100%/9). Otherwise, the probabilities are chosen so that the best-informed participant has twice the chance of receiving any information as the second-best informed participant (factor 2), and so on. Participants are chosen at random regarding their place in the knowledge distribution (i.e., if they are more or less well informed). • Decision rules used by the leader to make decisions • Consideration of hierarchy by the participants: Participants can weigh decisions higher or lower, depending on the position of the sender in the hierarchy relative to their own. • Presence of transactive memory within the group: If transactive memory is present, participants know if the sender is an expert in the information being sent and can weigh the input higher or lower (depending on their expertise).

			• Interaction pattern within the group
		I.ii.d If applicable, how is space included in the model?	Space is not included in the model.
		I.ii.e What are the temporal and spatial resolutions and extents of the model?	One time step is one discussion round, meaning that a participant is chosen to speak, the speaker provides a piece of information, and all other participants process it. See I.iii.a for the steps that happen during one simulated risk work-shop.

	I.iii Process overview and scheduling	I.iii.a What entity does what, and in what order?	 Stages of the simulation before, during, and after the risk workshop (adapted from Harten *et al.*, 2022).

V) Design Concepts	II.i Theoretical and Empirical Background	II.i.a Which general concepts, theories or hypotheses are underlying the model's design at the system level or at the level(s) of the submodel(s) (apart from the decision model)? What is the link to complexity and the purpose of the model?	The model has been developed to investigate the impact of choices made during the facilitation of a risk workshop, given a calculative culture of quantitative skepticism (Mikes, 2009). While some results are specific to decision-making regarding risk assessment, the general concept can be transferred to other settings where a group shares distributed knowledge in a discussion to make a decision. The interaction of the participants is built upon the idea of participants forming a transactive memory system (Wegner, 1987); all relevant information for a decision is available to the group, but the group needs to make use of the available information correctly. The experiments investigate deviations from an ideal speech situation (Habermas, 1982), as discussed in Harten *et al.* (2022).
		II.i.b On what assumptions is/are the agents' decision model(s) based?	The participants' decisions are based on heuristics that aim to reasonably account for all available information relevant to the respective decision. The participants are either following simple mathematical formulas for their decisions (e.g., when the leader determines if the discussion should continue), or make random choices, where model parameters determine the probability of each option to be chosen (e.g., when participants decide what information to share with the others).

		II.i.c Why is a/are certain decision model(s) chosen?	For some settings, there are no clear-cut rules available for how our participants will decide to act during the discussion. Thus, when participants need to make decisions, we allow participants to choose randomly between all possible options (e.g., when they decide who should be the next sender). However, the probability of each choice is influenced by reasonable heuristics (e.g., when the group is aware of the hierarchy, the facilitator should be more likely to choose agents with a high position in the hierarchy as the following senders).
		II.i.d If the model / a sub-model (e.g., the decision model) is based on empirical data, where does the data come from?	The model is not based on empirical data.
		II.i.e At which level of aggregation were the data available?	Not applicable.
	II.ii Individual Decision Making	II.ii.a What are the subjects and objects of decision-making? On which level of aggregation is decision-making modeled? Are multiple levels of decision making included?	Participants make two types of decisions: If they are chosen to be senders in a given discussion round. They choose what information to talk about. If they are chosen to be receivers, they decide how to weigh the input they receive from the sender. The facilitator makes one specific decision: In each round, one participant is chosen by the facilitator to be the next sender.

			The leader makes two specific decisions: After each round, the leader decides if the conditions for the decision and termination approaches are met, and therefore, a decision can be made. The leader also makes the final decision on how to assess the risk. The objective of the leader is to reach an accurate risk assessment in the shortest possible time. However, the criteria to end the discussion is not determined by the leader but is a model parameter. The leader might decide to end the discussion and make a decision if the group has not progressed for some time.
		II.ii.b What is the basic rationality behind agents' decision-making in the model? Do agents pursue an explicit objective or have other success criteria?	Participants make their best effort to gain a correct understanding of the risk in order to reach an accurate risk assessment. They have no hidden agendas or other individual objectives.
		II.ii.c How do agents make their decisions?	Participants make their decisions by random choice; however, the probability of each decision might not be equal. For example, participants are more likely to talk about information they are experts on.
		II.ii.d Do the agents adapt their behavior to changing endogenous and exogenous state variables? And if yes, how?	Agents do not adapt their behavior.

		II.ii.e Do social norms or cultural values play a role in the decision-making process?	The selection of the next sender is influenced by social norms (e.g., prioritize participants based on hierarchy) in some experimental settings.
		II.ii.f Do spatial aspects play a role in the decision process?	Not applicable.
		II.ii.g Do temporal aspects play a role in the decision process?	The leader might decide to end the discussion and make a decision if the group has not progressed for some time.
		II.ii.h To which extent and how is uncertainty included in the agents' decision rules?	The participants can take uncertainty regarding the correct risk assessment into account, e.g., when the facilitator chooses the next sender.
	II.iii Learning	II.iii.a Is individual learning included in the decision process? How do individuals change their decision rules over time as consequence of their experience?	The agents' understanding of the risk under assessment is modeled as an individual constraint satisfaction network for each agent. Participants update their understanding of the risk, based on the input they receive during the discussion from other participants. Their understanding of the risk determines their individual risk assessment. They do not learn beyond the discussion of an individual risk, i.e., there is no interaction spanning several simulated discussions.

		II.iii.b Is collective learning implemented in the model?	By exchanging their individual knowledge, the participants' understandings of the risk may move towards a more common understanding.
	II.iv Individual Sensing	II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? Is the sensing process erroneous?	Not applicable.
		II.iv.b What state variables of which other individuals can an individual perceive? Is the sensing process erroneous?	In some experimental conditions, participants know other participants' relative position in the hierarchy and other participants' expertise. This knowledge, if available, is free of errors. The facilitator and the leader know the overall risk assessment of all participants and can choose the next sender and decide when to end the discussion based on this knowledge.
		II.iv.c What is the spatial scale of sensing?	Not applicable.
		II.iv.d Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables?	Participants are assumed to know variables like the expertise of other participants (if the corresponding experimental condition is present). Participants learn about other participant's understanding of the risk via the simulated discussion.

		II.iv.e Are costs for cognition and costs for gathering information included in the model?	The cost of cognition and information gathering is accounted for by the fact that the leader makes a decision as soon as specific decision and termination criteria are met, instead of continuing the discussion potentially infinitely. However, the participants make no conscious decision on whether to invest cognitive resources: whenever they get new input, they update their knowledge about the risk and their risk assessment.
	II.v Individual Prediction	II.v.a Which data uses the agent to predict future conditions?	Participants do not predict future conditions.
		II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?	Not applicable.
		II.v.c Might agents be erroneous in the prediction process, and how is it implemented?	Not applicable.
	II.vi Interaction	II.vi.a Are interactions among agents and entities assumed as direct or indirect?	The participants directly interact with each other by exchanging information about the risk during the discussion.

		II.vi.b On what do the interactions depend?	In each discussion round, one participant is chosen to talk to all other participants. The other participants are only receivers in that discussion round but can become senders themselves in subsequent rounds.
		II.vi.c If the interactions involve communication, how are such communications represented?	Participants communicate by exchanging information about their individual knowledge about the risk. Usually, they will share the activation of their information node ("their knowledge") they have decided to share. The agents will deduce themselves how the new information relates to nodes already in their network. This process is assumed to be objective, i.e., all agents will agree if two nodes are in an explanatory or contradictory relationship.
		II.vi.d If a coordination network exists, how does it affect the agent behaviour? Is the structure of the network imposed or emergent?	The setting for the simulation is one discussion by all participants. Therefore, each participant can send information to all other participants, if chosen to be the sender.
	II.vii Collectives	II.vii.a Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Are these aggregations imposed by the modeller or do	The individual participants form a group. The risk assessments of the individuals can be aggregated to a decision, like a consensus, a majority vote, or an average vote (used to determine if the group opinion is moving over time). The leader decides to end the discussion based on such aggregates.

		they emerge during the simulation?	
		II.vii.b How are collectives represented?	The collective (that is, the group of all participants) has no agency by itself and is only a conceptual component of the model. All decisions are made by individual participants.
	II.viii Heterogeneity	II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?	The group participants differ in their initial knowledge and expertise. They might differ regarding their position in a hierarchy if the corresponding experimental condition is present. The participants have heterogeneous mental models initially, as the risk structure of their mental models depends on the information provided to them.
		II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?	Agents are not heterogeneous in their decision-making.
	II.ix Stochasticity	II.ix.a What processes (including initialization) are modeled by assuming they	For each run, a randomly generated risk is chosen for the group to assess.

		are random or partly random?	The decision by the facilitator of who is the next sender in a discussion round is random. Depending on experimental settings, the probability of each participant becoming the next sender can be unequal (e.g., when priority is given to some participants based on their risk assessment). The decision by the sender on what to share is random. The probability of each node being chosen for sharing can be unequal, depending on experimental settings.
	II.x Observation	II.x.a What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected?	At the beginning of each run, the independent variables are recorded. After each discussion round, the activation of all states of the 'overall assessment' node is recorded for each participant. Furthermore, it is checked and recorded if a decision and termination rule has been activated during the most recent discussion round and what decision by the leader it would result in.
		II.x.b What key results, outputs or characteristics of the model are emerging from the individuals? (Emergence)	We can observe how limitations to the ideal speech situation impact the effectiveness of risk workshops, measured as the share of high and low risks that get assessed correctly, as well as the time it takes to make a decision.
VI) Details	II.i Implementation Details	III.i.a How has the model been implemented?	The simulation is entirely written in Python. The original ECHO code was re-implemented in Python for this study.

		III.i.b Is the model accessible and if so where?	The model is available at CoMSES.
	III.ii Initialization	III.ii.a What is the initial state of the model world, i.e., at time t=0 of a simulation run?	<u>The benchmark assessment and the benchmark process</u> The simulation requires a benchmark assessment that the assessments reached in the risk workshop can be compared against. The benchmark assessment is determined by calculating a constraint satisfaction network with full information about the true risk. This is the assessment that an agent would reach if all information relevant to the risk assessment were available from the beginning. <u>Initial information distribution</u> In the beginning, information concerning the risk is distributed to the participants. After this initial seeding of (true) information, participants only get new input from other participants. All participants are able to assess the overall risk based on the limited knowledge they are pro-vided initially. Each information about the risk is available to at least one participant, so a correct risk assessment would be achieved if all participants could ideally share their information. The 27 pieces of information are distributed among the participants, either equally or unequally. <u>Knowledge about each other's hierarchical position</u>

			Each participant is assigned a specific hierarchical status (low, medium, high). If the corresponding experimental condition is present, participants are aware of the hierarchical status of the other participants and consider it when including sender input into their individual constraint satisfaction network. <u>Knowledge about each other's expertise</u> Information provided to the participants initially constitutes their expertise regarding the risk. Depending on the experimental condition, participants might be informed about each other's expertise, i.e., who is an expert concerning which information.
		III.ii.b Is initialization always the same, or is it allowed to vary among simulations?	Several simulation experiments are conducted with the model. Depending on the experiment, the following attributes are systematically varied: <u>Initial distribution of information among the risk workshop participants</u>: Either all participants get the same amount of information, or information is assigned unequally. Each information from the pool of information that is initially assigned to the group is distributed one after another. For each information, each participant has a specific probability of receiving this information. If information is distributed equally, each agent's probability of receiving a piece of specific information is 11.1% (i.e., 100%/9). Otherwise, the probabilities are chosen so that the best-informed participant has twice the chance of receiving any information as the sec-

			ond-best informed participant (factor 2), and so on. Participants are chosen at random regarding their place in the knowledge distribution (i.e., if they are more or less well informed). Consideration of hierarchy by the participants: Participants can weigh decisions higher or lower, depending on the position of the sender in the hierarchy relative to their own. Presence of transactive memory within the group: If transactive memory is present, participants know if the sender is an expert in the information being sent and can weigh the input higher or lower (depending on their expertise).
		III.ii.c Are the initial values chosen arbitrarily or based on data?	The initial values are chosen arbitrarily within the constraints set by the model parameter.
	III.iii Input Data	III.iii.a Does the model use input from external sources such as data files or other models to represent processes that change over time?	Not applicable.
	III.iv Submodels	III.iv.a What, in detail, are the submodels that represent the processes listed in	Sender selection The facilitator chooses the sender in a random draw that depends on the experimental condition. Only agents who have any expertise in any node are eligible

		'Process overview and scheduling'?	to become senders, as otherwise, the sender has no information to share. The following interaction pattern are investigated: Random: In the baseline model, the next participant to speak is chosen at random, with an equal probability for each participant. Priority is given to concern: The probability of each participant to be chosen is weighted by their overall risk assessment. Priority is given to dissent: The probability of each participant to be chosen is weighted by the distance of their individual risk assessment from the average group risk assessment. Priority is given to hierarchy: Participants are more likely to be the sender if they are assigned a higher hierarchical position. The probability of each participant to be chosen is weighted by *weight_h_low*, *weight_h_medium*, or *weight_h_high*. Priority is given to homogeneity: The probability of being the next sender is higher if the participant's risk assessment is close to the average group risk assessment. <u>Selection of sender output</u> The senders chose one of the information nodes available to them. The chance of choosing a specific node is weighted by the sender's expertise regarding the

			information node. The expertise is *weight_e_expert* for nodes initially assigned to the sender and *weight_e_nonexpert* for all other nodes. The sender will provide the detailed state of the information node to all other participants, who are the receivers in the given round. Furthermore, participants prioritize talking about nodes that have not been discussed often. The previously determined weight for each node ($node_i$) based on the participant's expertise regarding the node is multiplied by a weight reflecting how much it has been talked about by the group already, relative to the node that was talked about most often ($node_{max}$): $$weight(node_i) = weightExp(node_i) * 2^{discussionCounter(node_{max}) - discussionCounter(node_i)}$$ Here, *weightExp*($node_i$) is either *weight_e_expert* if the sender is an expert on $node_i$or *weight_e_nonexpert* otherwise. $discussionCounter(node_i)$ gives the number of times $nodei$ has already been discussed in this simulation run. <u>Inclusion of sender input</u> All receivers update their beliefs about the node the sender talks about by assigning the input a weight relative to their prior belief. Depending on the experimental condition, the weight assigned to the input reflects the following aspects: • difference in expertise regarding the specific node between sender and receiver;

			• difference in hierarchy between sender and receiver. For example, the receiver will weigh the input higher if the sender is an expert and the receiver is not, or if the sender has a higher position in the hierarchy than the receiver. For example, if a non-expert assessed the activation of an information node to as 0 and received the information of an expert that assesses the activation to be 0.23, the updated belief of the non-expert will be – $0.23 * (\text{weight\_e\_expert} - \text{weight\_e\_nonexpert}) - 0.0 * (1 - (\text{weight\_e\_expert} - \text{weight\_e\_nonexpert}))$
		III.iv.b What are the model parameters, their dimensions and reference values?	See Table 1.
		III.iv.c How were submodels designed or chosen, and how were they parameterized and then tested?	The submodels were designed to reflect a prototypical implementation of the processes during a risk workshop. The experimental conditions were derived by Harten *et al.* (2022) as deviations from an ideal speech situation and implemented as simply as possible. All decisions made by agents are either random or influenced by simple heuristics.

Parameters of the model

Parameter	Explanation	Default value(s)
nr_agents	Number of participants	9
threshold_change_continue	The maximum deviation in the group risk assessment for the discussion to be considered stagnant.	0.01
Risk model		
nr_issue_nodes	Number of children each parent-node has in the reference risk network	3
connectedness_issues	The probability that two issue nodes are connected by a constraint to each other	0.2
connectedness_information	The probability that an information node and an issue node are connected by a constraint to each other	0.2
connectedness_assessment	The probability that an issue node and a risk assessment node are connected by a constraint to each other	0.4
connections_rate_explanatory	The probability that a constraint is explanatory rather than contradictory	0.6
Experimental conditions		
information_distribution	Determines whether all agents have the same probability to receive an information in the initial distribution	equal/unequal
weight_hierarchy	Participants consider the difference in the hierarchy when weighting sender input	yes/no
weight_trans_mem	Participants consider the difference in expertise when weighting sender input	yes/no
mode_limited_trans	Participants consider their prior beliefs when integrating new information	yes/no
decision_rule	Decision-making rule used by the leader	
Interaction_pattern	Interaction pattern used in the group	

Appendix 6 Summaries

English summary

Due to the universally recognized importance of risk management for organizations, there are many guidelines and recommendations for good risk management practices. However, there is insufficient research on the actual impact of how risk management is implemented by the organization. This is especially true for the individual tools used for risk management, like risk workshops.

This thesis provides two studies that investigate how the design of risk workshops impacts the workshops' effectiveness. The first study uses an ideal speech situation as a starting point to derive deviations from the ideal that are commonly found in actual discussions. The impact of the deviations is evaluated by conducting agent-based simulation experiments using a computational model of a risk workshop. The results of the study provide recommendations for risk workshop facilitators on how their choices impact the workshop's effectiveness and identify trade-offs that have to be made based on the organization's risk attitude.

The second study investigates the importance of the calculative culture prevalent in an organization for the design of effective risk workshops. It builds upon the first study by using the same experimental framework but using a different cognitive architecture for the agents to reflect a different calculative culture. The study finds that the risk workshop facilitator needs to take the prevalent calculative culture into account when conducting a workshop, as the calculative culture changes the dynamics of the workshop and the effect of some deviations from an ideal discussion on the effectiveness of the workshop.

Besides providing guidance for workshop facilitators and a research framework to further investigate drivers of the effectiveness of risk workshops, the thesis contributes a novel approach to combine agent-based modeling of group work with complex cognitive architectures.

Deutsche Zusammenfassung

Da die Bedeutung des Risikomanagements für Unternehmen allgemein anerkannt ist, gibt es zahlreiche Leitlinien und Empfehlungen für gutes Risikomanagement. Es gibt jedoch noch zu wenige Erkenntnisse darüber, wie sich die Umsetzung des Risikomanagements in der Organisation tatsächlich auswirkt. Dies gilt insbesondere für die einzelnen Instrumente, die für das Risikomanagement eingesetzt werden, wie z. B. Risikoworkshops.

Für diese Arbeit werden zwei Studien durchgeführt, die untersuchen, wie sich die Gestaltung von Risikoworkshops auf die Effektivität der Workshops auswirkt. In der ersten Studie wird eine ideale Gesprächssituation als Ausgangspunkt genommen, um Abweichungen vom Ideal abzuleiten, die in realen Gesprächen häufig vorkommen. Die Auswirkungen dieser Abweichungen werden durch agentenbasierte Simulationsexperimente mit einem computergestützten Modell eines Risikoworkshops erhoben. Die Studie liefert Hinweise für die Moderatoren von Risikoworkshops, wie sich ihre Entscheidungen auf die Effektivität des Workshops auswirken, und zeigen auf, welche Kompromisse bei der Workshopgestaltung auf Basis der Risikoeinstellung des Unternehmens eingegangen werden müssen.

Die zweite Studie untersucht die Bedeutung der in einer Organisation vorherrschenden *Calculative Culture* für die Gestaltung effektiver Risikoworkshops. Sie baut auf der ersten Studie auf, indem sie denselben experimentellen Rahmen verwendet, aber eine andere kognitive Architektur für die Agenten nutzt, die einer anderen *Calculative Culture* entspricht. Die Studie zeigt, dass der Moderator eines Risikoworkshops die vorherrschende *Calculative Culture* bei der Durchführung eines Workshops berücksichtigen muss, da die *Calculative Culture* die Dynamik des Workshops und die Auswirkungen einiger Abweichungen von einer idealen Diskussion auf die Effektivität des Workshops beeinflusst.

Neben der Bereitstellung eines Leitfadens für Workshop-Moderatoren und eines Forschungsrahmens für die weitere Untersuchung der Faktoren, die die Effektivität von Risikoworkshops beeinflussen, liefert die Arbeit einen neuartigen Ansatz, der die agentenbasierte Modellierung von Gruppenarbeit mit komplexen kognitiven Architekturen kombiniert.

Appendix 7 List of publications

As part of this dissertation project, the following publications have been published:

- Bellora-Bienengräber, L., Harten, C. and Meyer, M. (2023), “The effectiveness of risk assessments in risk workshops: the role of calculative cultures”, *Journal of Risk Research*, Routledge, Vol. 26 No. 2, pp. 163–183.
- Harten, C. (2019), “Agent-Based Model of Risk Assessment: A Distributed Cognition Approach”, in Linsley, P., Shrives, P. and Wieczorek-Kosmala, M. (Eds.), Multiple Perspectives in Risk and Risk Management, Springer International Publishing, pp. 169–178.
- Harten, C., Bellora-Bienengräber, L. and Meyer, M. (2023), “Effektive Risikoworkshops … Aber wie?”, *Controlling*, Vol. 35 No. S, pp. 50–54.
- Harten, C., Meyer, M. and Bellora-Bienengräber, L. (2022), “Talking about the likelihood of risks: an agent-based simulation of discussion processes in risk workshops”, *Journal of Accounting & Organizational Change*, Vol. 18 No. 1, pp. 153–173.

Erklärung

Hiermit erkläre ich, Clemens Harten, dass ich keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die Arbeit wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

_______Hamburg, 30.07.2025____ ______________________________
Ort/Datum Unterschrift Doktorand/in

****

Eidesstattliche Versicherung:

Ich, Clemens Harten, versichere an Eides statt, dass ich die Dissertation mit dem Titel:

„ Effectiveness of risk workshops: an agent-based simulation study“

selbst und bei einer Zusammenarbeit mit anderen Wissenschaftlerinnen oder Wissenschaftlern gemäß den beigefügten Darlegungen nach § 6 Abs. 3 der Promotionsordnung der Fakultät für Wirtschafts- und Sozialwissenschaften vom 18. Januar 2017 verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht benutzt.

_______Hamburg, 30.07.2025____ ____________________________________
Ort/Datum Unterschrift Doktorand/in