safety

MDPI

Article

Can Complexity-Thinking Methods Contribute to Improving Occupational Safety in Industry 4.0? A Review of Safety Analysis Methods and Their Concepts

Arie Adriaensen [1,*], Wilm Decré [2,3] and Liliane Pintelon [1]

[1] Centre for Industrial Management/Traffic and Infrastructure, KU Leuven, Celestijnenlaan 300 – box 2422, 3001 Leuven, Belgium; liliane.pintelon@kuleuven.be

[2] Robotics Research Group, KU Leuven, Celestijnenlaan 300 – box 2420, 3001 Leuven, Belgium; wilm.decre@kuleuven.be

[3] Flanders Make@KU Leuven, Core Lab ROB, Celestijnenlaan 300 – box 2420, 3001 Leuven, Belgium

* Correspondence: arie.adriaensen@kuleuven.be

Received: 24 August 2019; Accepted: 27 September 2019; Published: 3 October 2019

Abstract: With the introduction of Industry 4.0, occupational health and safety finds itself confronted with new types of hazards. Many Industry 4.0 innovations involve increased machine intelligence. These properties make socio-technical work in Industry 4.0 applications inherently more complex. At the same time, system failure can become more opaque to its users. This paper reviews and assesses safety analysis methods as the breakdown of interaction coupling in socio-technical systems on the one hand, and the degree of failure tractability on the other hand; the latter being used as a proxy for complexity. Previous literature confirms that traditional health and safety risk assessment methods are unable or are 'ill-equipped' to deal with these system properties. This paper studies the need to introduce new paradigms and safety methods related to complexity thinking with theories borrowed from the study of complex adaptive systems, all to assess the arena of abruptly changing hazards introduced by Industry 4.0. At the same time, this review makes clear that there is no one-solution-fits-all method. Occupational health and safety (OHS) covers many different hazard types and will need a combination of old, new and yet-to-be-developed safety assessment methods.

Keywords: industry 4.0; occupational safety; complexity thinking; systems thinking; safety-II; safety analysis methods; resilience engineering; risk assessment

1. Introduction

The understanding of work systems in industrial safety has fundamentally changed with the introduction of Industry 4.0. This type of industry (so named after the fourth industrial revolution) is the successor of the three previous industrial revolutions. The first industrial revolution in the 18th century commenced with mechanization, the second introduced the use of electricity and the third revolution adapted digitalization [1]. The fourth industrial revolution is mainly about interconnection via the internet of things (IoT) [2], advances in smart use of digitized information and technology, responsive devices in so-called 'smart factories' and industrial automation via autonomous algorithms and artificial intelligence (AI) [3]. Industry 4.0 not only entails new technologies, but also transforms traditional work systems into highly interconnected networks of people, technology and business units. It is also marked by increasing competitive pressure to implement mass customization solutions [4].

Although the academic world is still debating a uniform definition for Industry 4.0 [5,6], it can be observed that Industry 4.0 innovations currently are mainly introduced in the manufacturing industry [2]. Other areas include logistics and processing industries, including agri-food and construction [7,8]. Such industries apply risk analysis methods, many of which were developed more than half a century ago. Examples are governing industrial safety analysis methods [9] such as root cause analysis (RCA), developed before World War II [10]; failure mode and effects analysis (FMEA), developed in 1949; and hazard and operability study (HAZOP), developed in the 1960s [11]. Such methods were designed with a certain type of risk in relation to the relatively stable conditions of the past. Meanwhile, contemporary work systems have become more complex and are characterized by new emerging risks (NER) [12,13]. One critique that applies to traditional safety analysis methods is that they are only able to analyze linear cause–effects, and that they cannot solve complex or emergent system behavior [10,14–16].

New emerging risks originate from "new processes, new technologies, new types of workplace, or social or organizational change", but also because "new scientific knowledge allows a long-standing issue to be identified as a risk" [12] (p. 7). One example of this is the threat to occupational health from new materials that require novel safety analysis methods [17]. The European agency for safety and health at work warns in its expert forecast on emerging physical risks [18] of the increase and the inappropriate design of complex human–machine interaction. The report calls for adapting a methodology for identifying errors in the design of human–machine interactions, the use of proper feedback design and the adequacy of information provided to the user via training materials [18]. In relation to NER [19] 'complexity of human–machine interfaces' is mentioned as a specific ergonomic challenge. The expert forecast on NER from the EU Agency for Safety and Health at work [18] identifies that complexity of new technologies with their resulting transformation of work processes, and poor design of human–machine interfaces, lead to increased mental and emotional strain on workers. Others have described that although automation in many circumstances yields better system performance, it also comes at the cost of producing more problematic performance when things fail in comparison with non-automated systems [20]. This is referred to as the lumberjack effect, an analogy to trees in the forest and how "the higher they are, the farther they fall" [20] (p. 477). This was already true in earlier industrial revolutions where one study showed that assembly line automation in the Norwegian furniture industry reduced the cutting of fingers, but maintenance and handling of disruptions resulted in more severe injuries, such as the amputation of arms [21].

This paper takes complexity challenges such as interconnectivity, autonomous systems, automation in joint human–agent activity and a shift in supervisory control as essential traits in Industry 4.0 (see Section 2). Increased machine autonomy, the use of collaborative robots and wide-scale interconnectivity between software, hardware and artefacts both within and across business units urges the safety community to reflect on new challenges for occupational health and safety (OHS) in Industry 4.0. They cause a shift in the nature of safety-critical sociotechnical systems and expose "theoretical and methodological flaws in contemporary accident analysis methods" [22] (p. 164) and causation models. The safety of the interconnectivity challenge is not solved by simply adding more reliability to interfaces, but transforms conventional work practices, adds complexity and can reduce transparency of underlying processes.

This study will introduce safety analysis principles that emerged under the paradigms and concepts of (i) systems thinking, which takes a holistic perspective to systems as a whole; (ii) resilience engineering (RE), as the response to external and internal stressors and their recoveries from breakdowns and; (iii) Safety-II thinking, which marks a shift from learning from failure to learning from the full performance variability of a system. Ultimately these concepts are all related to complexity thinking with its focus on emerging behavior from complex dynamic interactions. These concepts will be explained in detail in Sections 3.2.1–3.2.4.

Badri [13] made a comprehensive literature review of the status of OHS in Industry 4.0 and came to the conclusion that only a small number of publications on the subject exist. The review also concluded that research on technical advances in Industry 4.0 rarely cited the integration of OHS, and observed that OHS initiatives and technological developments of different manufacturers in Industry 4.0 are fragmented. These aspects urge researchers, field experts and industrialists to collaboratively ensure a safe transition to this new paradigm [13]. In addition to the challenge of adapting technical and safety standards [5,13], “future OHS integration initiatives must combine at the outset virtual task analysis, dynamic evaluation of occupational risks, cognitive analysis of workload, and skills management tools” [13] (p. 409). These observations were a motivation for this paper, whereby we will present opportunities for Industry 4.0 borrowed from the safety analysis of complex adaptive socio-technical systems that ultimately led to the scope and formulation of our research question: “Can complexity thinking contribute to progress in occupational safety in Industry 4.0?” This paper will highlight some of the limitations of current OHS safety analysis methods for an Industry 4.0 environment, and present some of the strengths of complexity-thinking safety analysis methods with which we strive to expand the assessment method assortment for industrial safety management. Such methods will be introduced to the reader. It will be examined whether these complexity-thinking methods match the challenges specific to Industry 4.0. This extends the work from former studies that have examined a similar shift in relation to accident investigation methods [14], or studies that cover in more detail a tendency towards systems and complexity thinking in safety management in general [10,16,21–25]; this is in answer to the changing nature of industrial risk [11,13,21].

The remainder of the paper is organized as follows: Section 2 describes the methodology for this study; Section 3 describes the results, starting with an analysis of traditional safety methods, an exploration of contemporary safety paradigms and their principles and applicable methods from outside Industry 4.0, as well as how they answer exposed challenges and shortcomings; Section 4 discusses and draws conclusions from these results.

2. Materials and Methods

The research was performed by a desktop exercise that analyzes existing and novel safety methods and how these methods answer the challenges of NER in Industry 4.0. Figure 1 depicts how ‘The Answerability to Industry 4.0 Problem Space’ (green) results from the convergence of the examination of ‘Answerability Safety Methods’ (blue), on the one hand and exploration of the ‘Risks in Industry 4.0’ (yellow) on the other hand.

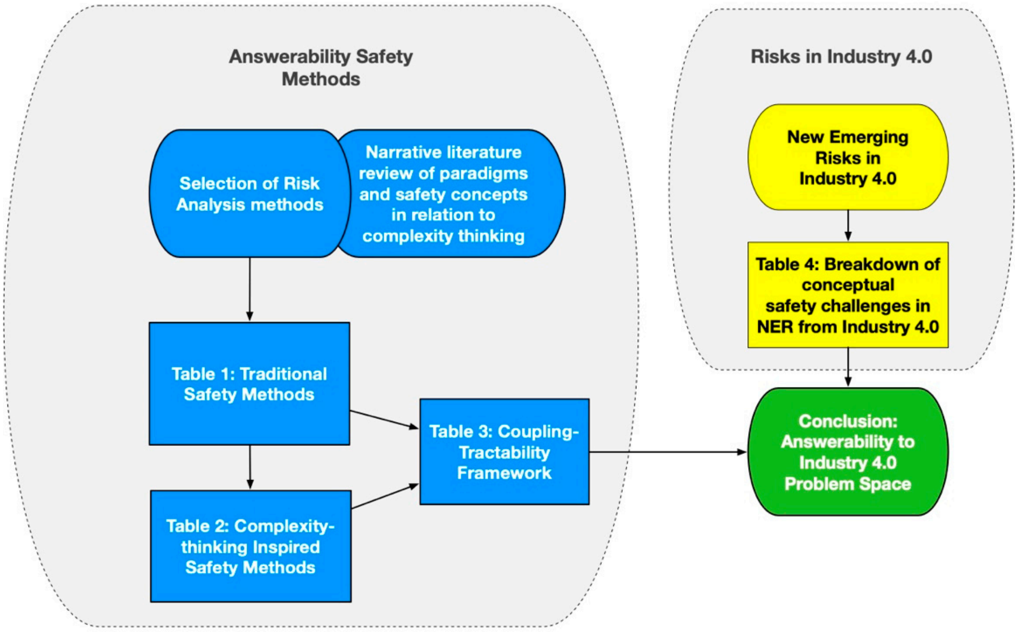
Answerability Safety Methods
Selection of Risk Analysis methods
Narrative literature review of paradigms and safety concepts in relation to complexity thinking
Table 1: Traditional Safety Methods
Table 2: Complexity-thinking Inspired Safety Methods
Table 3: Coupling-Tractability Framework
Risks in Industry 4.0
New Emerging Risks in Industry 4.0
Table 4: Breakdown of conceptual safety challenges in NER from Industry 4.0
Conclusion: Answerability to Industry 4.0 Problem Space

Figure 1. Methodology.

The analysis under ‘Answerability Safety Methods’, focuses on a literature study about safety analysis methods from traditional safety methods and developments in contemporary safety science related to complexity thinking. First of all, a list of traditional safety methods in relation to occupational safety was gathered on the basis of the safety analysis methods contained in two sources: (i) *Safety*

Analysis, Principles and Practice in Occupational Safety from Harms-Ringdahl [9], with a comprehensive review of methods and; (ii) *ISO 31010—Risk Management—Risk Assessment Techniques* [26]. We checked additional safety analysis method compilations from Lees (1980) [27], Johnson (1980) [28], ILO (1988) [29] and Bahr (1997) [30]. These compilations were suggested by Khanzode et al. [31] to be the main references together with the Harms-Ringdahl publication. ISO 31010 was an additional contribution not mentioned by Khanzode et al. ISO 45001, with title, *Occupational Health and Safety Management Systems* [32], was additionally consulted and, although it is very relevant for OHS, could be considered to act on a meta level of measuring and mitigating safety. It displays the requirements and methods to manage safety systems on a safety-management level, but does not list safety analysis methods of the work-system level like ISO 31010. Notwithstanding its relevance as an integrated OHS management approach, it was considered to be outside the scope of collecting and reviewing safety analysis methods.

The references from Harms-Ringdahl and ISO 31010 selected in this study are the most recent sources from our literature search. In combination, they produce the broadest collection of methods, whereas the remaining sources mainly resulted in a duplication of methods, when management and intervention techniques were excluded. Management and intervention techniques were the main focus in the ILO [29] and Bahr [30] references. The focus of this study is on causation and understanding risks, not on understanding intervention, injury models or severity of consequences, for which additional info can be found in Khanzode et al. [31]. Overlap of methods between the two selected sources was resolved and a single list of methods was produced. Whereas ISO 31010 is the principal document to consult for many safety managers in an OHS environment, we began with Harms-Ringdahl's academic book as the primary source as this author provided a collection of well-defined selection criteria, which were based on the following principles: (i) support for a systematic approach; (ii) ease of understanding and application; (iii) possibility of being able to apply analysis even when information is incomplete; and (iv) taking into consideration that the analysis can be performed with a reasonable amount of effort [9]. This produced a primary set of 10 methods, as pre-selected by Harms-Ringdahl. Thereafter, we classified remaining safety analysis methods according to their paradigms. For each paradigm that was not yet covered by the primary set of 10 methods according to the selection criteria of Harms-Ringdahl, we retained one method to complement the list. Where supplementary reasons justified the further inclusion of methods, although previously covered by a method from the same paradigm, we provided additional rationale for inclusion. This process resulted in the collection of methods from Table 1 (presented in Section 3.1). In this table, each method was further assigned with (i) a short description; (ii) its underlying paradigm; (iii) its unit of analysis; and (iv) its capabilities to analyze couplings and complexity interactions (the latter expressed as tractability).

In a subsequent step, the literature was examined for progress in safety science about how complexity-thinking-related paradigms exposed shortcomings in many traditional safety thinking methods from Table 1. We used literature gathered from previous research projects advocating the use of these new concepts and paradigms. We additionally examined the scientific databases Scopus and Web of Science for literature that challenged such concepts, to produce an objective narrative literature review. Although very few challenges to the complexity-thinking concepts were found, we were able to complement our existing set of papers in favor of the approaches we identified during our search of these databases. The results can be found in Section 3 in the form of descriptions of systems thinking, complexity thinking, Safety-II and resilience engineering along with an explanation of how these schools of thought responded with new theoretical underpinnings and concepts. Subsequently, the literature was analyzed for complexity-thinking-related safety methods that could overcome some of these shortcomings. The results can be found in Table 2 (presented in Section 3.2.6), which follows the same template as Table 1 for the traditional safety methods. Although the sources from the traditional methods in Table 1 used an assortment of methods, models, frameworks and tools, we restricted ourselves to the collection of methods only when identifying complexity-inspired safety analysis methods as candidates for assessing Industry 4.0 challenges. Methods are accompanied by structured instructions for systematic hazard identification and/or risk assessment, whereas frameworks and models lack a logical and structured approach to achieve validation and verification of their outcomes.

The last columns from Table 1 and Table 2 were used to produce a matrix with four quadrants to position the safety analysis methods in terms of the earlier-derived dimensions of coupling and tractability. This matrix is an adaption of a framework introduced by Perrow (1984) and further adapted by Hollnagel et al. (2008). The framework is described in Section 3.1 as part of the explanations on complexity thinking. The results from Tables 1 and 2 and the resulting Table 3 (presented in Section 3.2.6 emerged from expert judgement from a panel that consisted of the lead author and an external mechanical and aerospace engineering researcher, with expertise on the analysis of complex socio-technical systems, mainly in terms of risk and safety management, and resilience engineering. The expert judgement was supported with references to the literature. Where needed, differences in opinions were resolved by providing additional rationale from multiple literature sources. The results were verified by the co-authors, being mechanical engineering and industrial management researchers with experience in safety science and previous Industry 4.0 involvement.

The yellow-colored 'Risks in Industry 4.0' (the second segment in the methodology, as depicted in Figure 1) focuses on NER in Industry 4.0. The starting point is a collection of what Badri et al. [13] summarize in their introduction as a set of non-negligible assets in Industry 4.0: "Real-time communication, big data, man-machine cooperation, remote sensing, monitoring and control, autonomous equipment and interconnectivity" [13] (p. 403). From this list of new Industry 4.0 challenges, we restricted our scope to human operator-related issues in a socio-technical context and we consequently excluded 'real-time communication' and 'big data'. Although these challenges are—in our opinion—equally important, we believe that their primary technological nature, notwithstanding their socio-technical influence, deserves an independent analysis in separate research. Given that the effects from real-time communication in previous safety fields produced substantially less search returns in scientific databases, the research effort is comparatively large, and we restricted ourselves to the topics previously described. The subject is also primarily inversely related to safety. It is the 'lack' of real-time communication that poses a threat to the use of new technologies in Industry 4.0. In relation to Industry 4.0, the fragmentation of big data methodologies lay between process and maintenance departments and the need for prognosis instead of detection and diagnosis are mentioned as future challenges [33], but are outside our scope.

We restricted our scope to accident causation in terms of hazard identification and risk assessment, and did not take injury models into account; this deserves an analysis of its own and forms a different discipline. Thereafter, we translated the four remaining principles from the Badri et al. [13] introduction into four key challenges: interconnectivity, machine autonomy, automation in joint human–agent activity and a shift in supervisory control, the latter being derived from remote sensing, monitoring and control. These principles are also reoccurring design principles and technological features in the Industry 4.0 literature (e.g., [1,5]), whereby in Herman et al. [5] a shift in supervisory control is mainly represented as the aspect of decentralized decision making. Subsequently, we matched theoretical concerns from previous safety literature to these four key challenges. We consulted a cognitive systems engineer researcher for the identification of relevant papers regarding theoretical safety concerns. This related to our four overarching principles, to involve an additional relevance check for our literature search, and the process resulted in a set of four papers that produced a breakdown of the four main principles as the result of an exploratory process and without claim that we have identified an exhaustive breakdown (nor can such a process cover all scientific opinions). The results can be found in Table 4 and are further explained in Section 3.3. The capabilities of the methods in Tables 1 and 2 in relation to the safety challenges from Table 4 (presented in Section 3.3) will be explored as an answer to the research question: "Can complexity thinking contribute to progress in occupational safety in Industry 4.0?".

3. Results

The results confirm the line of thought set out in the methodology, starting from the analysis of traditional safety methods as used in OHS, the advances in theoretical developments of safety science and how these developments meet the challenges of Industry 4.0.

3.1. Safety Analysis Methods

Table 1 displays a set of traditional occupational safety methods taken from the two authoritative sources [9,26], from which Harms-Ringdahl's set of 10 primary methods forms the core of the table.

Table 1. Safety analysis methods (concept, paradigm and basis for structuring).

Method	Concept [9–26]	Paradigm	Basis for Structuring	Coupling/Tractability
Energy Analysis [9]	Identifies energies that can harm human beings.	Energy barrier thinking	Volumes that jointly cover the entire object	Loose coupling—tractable
Hazard and Operability Studies (HAZOP) [9–26]	Identifies deviations from intended design of equipment, based on the use of predetermined guide words. It is generally carried out by a multi-disciplinary team during a set of meetings.	Linear causality	Deviation of operational parameters	Tight coupling—tractable
Failure Mode and Effect Analysis (FMEA) [9–26]	Identifies failures of components and their effects on the system.	Linear causality and decompositional analysis	Reliability from components or modules	Tight coupling—tractable
Fault Tree Analysis [9–26]	Causal factors are deductively identified, organized in a logical manner and represented pictorially in a tree diagram that depicts causal factors and their logical relationships to the top event.	Energy barrier thinking, linear causality and decompositional analysis	Fault propagation resulting from initial event	Loose coupling—tractable
Event (Effect) Tree Analysis [9–26]	Analyzes alternative consequences of a specified hazardous event.	Energy barrier thinking, single cause philosophy and decompositional analysis	Fault propagation back to initial event	Loose coupling—tractable
Action Error Method [9]	Identifies departures from specified job procedures that can lead to hazards.	Taylorism	Phases of work of operator	Loose coupling—tractable
Job Safety Analysis [9]	Identifies hazards in job procedures.	Rationalist, prescriptive and top-down belief in procedures and Taylorism	Elements in an individual job task	Loose coupling—tractable
Deviation Analysis [9]	Identifies deviations from the planned and normal production processes.	Rationalist, prescriptive and top-down belief in procedures and Taylorism	Activities (e.g., activity flow or job procedure)	Loose coupling—tractable
Safety Function Analysis [9]	A structured description of a system's safety functions, including an evaluation of their adequacies and weaknesses.	Energy barrier thinking	Defenses or safety functions of the system	Loose coupling—tractable
Change Analysis [9]	Establishes the causes of problems through comparisons with problem-free situations.	Failure without acknowledging context sensitivity or emergent behavior	Discrepancy between as-is and as-should-be situation	Loose coupling—tractable
Root Cause Analysis (RCA) [26]	Attempts to identify the roots or original causes instead of dealing only with immediately obvious symptoms.	Single cause philosophy, linear causality and decompositional analysis	Initiating failure causes and effects	Loose coupling—tractable

Table 1. *Cont.*

Method	Concept [9–26]	Paradigm	Basis for Structuring	Coupling/Tractability
Human Reliability Assessment (HRA) [9–26]	Identification and prediction of human errors in relation to strictly predefined tasks.	Human reliability assessment and Taylorism	Human error	Loose coupling—tractable
Deterministic Probabilistic Risk Assessment (e.g., Risk Indices -FN Curves (the cumulative frequency 'F' of people affected 'N') [26]	Deterministic probabilistic risk assessment (PRA) produces a semi-quantitative measurement of risks based on frequency and severity scales.	(Semi-)quantitative causality credo	Ordinal or cumulative frequency and/or severity of harmful events	Loose coupling—tractable
Databases (e.g., Reaction Matrix—Consequence Analysis) [26]	Analysis of consequences of chemical risks like fire, explosions, the release of toxic gases or the determination of toxic effects or combinations of chemicals.	Database	Chemical and physical reactions	Loose coupling—tractable
Cognitive Task Analysis [26]	An analysis method that addresses the underlying mental processes that give rise to errors.	Task analysis as the key to understanding system mismatches	Tasks	Loose coupling—tractable
Bayesian Networks [26]	A method that use a graphical model to represent a set of variables and their probabilistic relationships. The network is comprised of nodes that represent a random variable and arrows that link parent nodes to a child nodes.	Epidemiological causation model	Events (and their related degrees of belief)	Tight coupling—tractable
Layer Protection Analysis (LOPA) [26]	LOPA is a semi-quantitative method for estimating the risks associated with undesired events or scenarios and the presence of sufficient measures to control them. A cause–consequence pair is selected, and the preventive layers of protection are identified.	Epidemiological causation model and energy barrier thinking	Multiple defenses	Tight coupling—tractable
Bowtie Analysis [26]	A simple, diagrammatic way of describing and analyzing the pathways of a risk from causes to consequences. The focus of the bowtie is on the barriers between the causes and the risk, and the risks and consequences.	Epidemiological causation model and energy barrier thinking	Multiple causes and defenses	Tight coupling—tractable

[9] Harms-Ringdahl, L.—Safety Analysis, Principles and Practice in Occupational Safety; [26] ISO31010—Risk Management-Risk Assessment Techniques.

The authors from this study recognized from their research involvement in OHS activities that these 10 methods are indeed among the most commonly used OHS methods, and many of them are additionally mentioned in ISO 31010. In Step 2 we labeled the remaining methods in line with the columns of Table 1 and gathered a total of 47 methods. To complement the initial set, we looked for methods from which the paradigm was not yet covered by another method. Before performing this task, we narrowed down the selection by disqualifying methods that lacked analytical structuring, and which we regarded as only have benefit as supporting tools for stand-alone methods. Examples of such disqualified tools, although mentioned by ISO 31010 as risk assessment techniques, include brainstorming, semi-structured interviews, the Delphi technique and multi-criteria decision analysis. However, some of these tools can lend powerful support from subject matter experts or stakeholder input for data-gathering techniques in hazard identification and risk assessment. They usually require an underlying causation model to structure the types of hazards or risks a researcher or safety manager is interested in. The analytic hierarchy process (AHP) is an example of a multi-criteria decision analysis with a strong analytical structuring that can be used as a hierarchy decision model for assessing the priority of OHS goals [34,35] (see also Section 4.2 for the use of AHP in combination with complexity-thinking methods). Consequentialist approaches that do not provide a comprehensive risk analysis with an objective basis for structuring hazards or risks like cost–benefit analysis or business impact analysis were removed. We also deleted a method that assessed safety culture and was not recognized as a primary means to achieve hazard reduction [36–38], but which belonged to a different categorical status of risk reduction by management technique.

3.1.1. Critical Analysis of Traditional Safety Methods

A brief explanation of the concepts of the methods can be found in Table 1. For a more extensive explanation, we refer to [9,26]. Table 1 reveals that the basis for structuring for a given method (the unit of analysis) varies considerably. The methods from Table 1 range from structuring the analysis by (i) failure propagation of events (fault tree analysis, event tree analysis, root cause analysis, Bowtie analysis); (ii) the defenses that protect against such failure propagations (safety barrier diagram, layer of protection analysis [LOPA]); (iii) deviations (HAZOP, what-if analysis, deviation analysis); or (iv) effects from component failure (FMEA) as the basic structuring unit. The assortment of units of analysis varies so greatly that any comparison between the methods becomes obsolete. This is an important reason why we have decided to follow the term 'safety analysis methods' throughout this paper for its general meaning. Safety analysis methods are shaped by a multi-step process consisting of (i) hazard identification; (ii) risk assessment by evaluation or prioritization, and subsequently (iii) mitigation and management of risks. In occupational safety, an additional step would be to produce injury mechanism models [31]. However, the shift in risks in Industry 4.0 is mainly a systems ergonomics challenge, an issue of organizing and distributing information and coordination. We hypothesize that similar physical forces will produce similar types of injuries, no different than before the introduction of Industry 4.0. Collaborative robots bring one exception, something that would indeed require new injury mechanism models because of the close interaction between workers and robots. We considered injury models to fall outside the scope of this paper.

From studying Table 1, one can learn that some safety analysis methods focus solely on hazard identification (energy analysis, primary hazard analysis, HAZOP) whereas others solely focus on risk assessment (risk indices, FN curves [whereby F stands for cumulative frequency and N for the number of people affected]), or a combination thereof (FMEA). Yet other methods only concentrate on the alleged sources for hazards, such as not following job instructions and procedures (job safety analysis, deviation analysis) or cognitive mismatches between operators and tasks, including the use of interfaces and system interactions (cognitive task analysis). The four challenges of interconnectivity, machine autonomy, automation in joint human–agent activity and a shift in supervisory control are all socio-technical-related issues. Most if not all traditional safety analysis methods need a combination of different methods to cover technical and human performance-related hazards and risks. There is a legitimate risk that analyzing such issues in isolation from each other says little or nothing about their integration on a higher-order level of human–machine interaction. Bowtie analysis would be an exception as it can cover a broad range of technical, human and systems-related issues. In Bowtie analysis, adverse events are depicted as central nodes. Such adverse events can be triggered by multiple contributing threats in combination, whereas none of these events in isolation would be sufficient to trigger the adverse outcome. Even the consequences could be multiple, with different effects on different parts of the socio-technical system. A framework on the 'joint integration' of technical and human performance safety issues via a combination of multiple methods is largely absent in *ISO 31010—Risk Management—Risk Assessment Techniques*, a leading reference for industrial safety managers. ISO31010 gives the user general guidance by tabling strength and weakness overviews with a general explanation for each method; however, it expects the user to self-assess which method or methods to use. Pasman et al. [39] propose a structured six-step approach, each with several methods tied to individual steps of the process: (i) hazard identification; (ii) quantification of consequence; (iii) quantification of probability of events; (iv) quantified risk; (v) risk reduction; and (vi) risk assessment. This multi-phased structured approach merges to some extent the non-commensurability of the units of analysis among the different methods. It can be hypothesized that such structured approaches in the academic literature will not always be consulted by safety managers in the field, where choice of method is influenced not only by several factors such as skills, experience capacity and capability of the risk assessment team, budget and resource availabilities [26], but also by industry and individual preferences. Pasman et al. [39] have provided some criteria for sound scientific risk assessment, yet warn that retrospective understanding of failure is much easier to achieve than prospective understanding of risks and hazards. In retrospect, "the cognitive illusion of 'What-You-See-Is-All-There-Is' (WYSIATI) gives rise to believing a one-sided explanation of an event, a confirmation bias, or the related anchoring effect and the availability bias" [15] (p. 81), whereas the 'What-You-Look-For-Is-What-You-Find' (WYLFIWYF) principle [40] means that the causes found during an analysis reflect the assumptions of the model used. One assumption that is valid for many occupational safety methods is that they are based on linear causal models and are therefore only able to account for linear cause–effects [10,14–16]. As such, these methods fail to explain non-linear or emergent system behaviors [14], something that is intertwined with the analysis of our Industry 4.0 challenges as introduced in Sections 1 and 2. Linear thinking is "a process that follows a chain of causal reasoning from a premise to a single outcome. In contrast, systems thinking regards an outcome as emerging from a complex network of causal interactions, and, therefore, not the result of a single factor" [23] (p. 939). Industry 4.0 is characterized by interconnectedness and autonomous algorithms. In such dynamic and highly coupled systems "[a]symmetry or non-linearity means that an infinitesimal change in starting conditions can lead to huge differences later on" [23] (p. 942) that arise from unusual combinations of conditions rather than from linear propagation of cause–effect chains [11].

3.1.2. An Historical Introduction to Traditional Causation Models

Several authors have described the historical evolution of accident and incident investigation methods and causation models [10,15], which started with linear models like Heinrich's Domino Model from 1931 [41], followed by Gibson's and Haddon's model of energy barriers [21] in the 1960s. Both causation models use "a closed system safety mindset with mechanistic metaphors to describe the conditions, barriers and linear chains of an accident process" [21] (p. 951). The Domino Model views accidents as event chains in a five-stage linear model, depicted as a line of dominoes. In order for the last two dominoes—the accident and resulting injury—to fall, Heinrich believed it was important for management to remove any of the first three dominoes, those being: (i) social factors (e.g., inheritance, environment); (ii) faults of people (e.g., violent temper, ignorance of safe practice); and (iii) the actual unsafe acts [40]. The analogy was that if any of these dominoes were to be removed, this would prevent the accident and injury dominoes from falling. Although the negative effects of an over-extensive linear extrapolation of protection measures have been highlighted by theories of error [42], most occupational safety causation methods are still based on the Domino Model and energy-barrier models [21] (see Section 3.2.4 for energy-barrier thinking in more complex environments). This is indeed confirmed by the findings in Table 1, whereby the many labels in the paradigm column are either based on linear causality, energy-barrier thinking or a combination thereof. Methods based on a single cause philosophy like root cause analysis or fault tree analysis depict a linear and deterministic cause–effect relation between consecutive events [43], and are therefore essentially another example of linear causality models.

3.1.3. Reflections on Quantitative Methods and Techniques

A last category is formed by quantitative methods which, although they have a common denominator, do not necessarily belong to a collective ontological paradigm. Despite their lack of an internal causation model, these methods provide helpful solutions that differ from those of previously discussed methods. We have therefore admitted quantitative methods and provided some supplementary reflections. Sneak circuit analysis and Markov analysis are engineering methods, although there are some rare occasions where Markov analysis has been used to predict human–machine interaction or socio-technical-related problems [44]. Their use in socio-technical systems and application in workplace safety has been, however, limited to simple systems, and their predictive utility remains limited to micro-level ergonomics [45]. Nonetheless, they should not be dismissed, as they remain important building blocks in understanding the micro-level system design of a larger chain of socio-technical challenges. Other probabilistic methods like Monte Carlo simulation and Bayesian networks assign probability to risks derived from a previous step of hazard identification. Monte Carlo simulations introduce a set of pragmatic solutions to model the total effect of uncertainty and can consequently handle a degree of intractability in its assessment. The accident causation representation in Bayesian networks is causal [46] and hence tractable, and the degree of unknown conditions inherently affects the quality of the model. On the other hand, by using conditional probability tables, more complex relationships can be modeled than with logical operators [46], and the application of Bayesian networks is not restricted to binary factors only. It can therefore be used to examine tight couplings. Bayesian networks have been used in the past to quantify the probability of human performance, in general to assess human error. Monte Carlo simulation has even been used to assign probabilities of dependencies [47,48] and their effects in conjunction with complexity-thinking methods. Although they can also be used to analyze tractable, loosely coupled failures, we will position Bayesian networks and Monte Carlo simulation in relation to their maximum analytical capacities along the quadrants of Table 3. It should be considered that in probabilistic risk assessment (PRA), the analyst is the one that discovers the risk scenarios, not the methodology [49]. PRA's predictive power is therefore limited. The history of application of PRA in NASA is a good example of the challenges related to its use. Early PRA models on the Apollo program, for example, were based on largely overestimated failure probabilities, which resulted in unrealistically low chances of mission success, whereas in the space shuttle era NASA seriously underestimated the rocket booster failure that caused the Challenger crash [50]. After the Challenger crash, NASA moved to adopt PRA from the nuclear industry [51], and succeeded in achieving quite accurate predictions during the last decades of the space shuttle program [49]. Currently, the combination of qualitative and quantitative methods is promising if applied to the proper scope of analysis. Aviation and space missions, for example, rely heavily on combinations of qualitative methods like FMEA, which is supported by quantitative approaches [52]. The future adoption of fuzzy logic into safety analysis, as a generalized theory to manage uncertainty ([53], or turn uncertainty in probability, is a remaining challenge for future research).

Highly deterministic probabilistic methods such as risk indices or FN curves (whereby F stands for frequency and N for the number of people affected) with binary logic and cause–effect constructs are not able to find anything else than tractable, loosely coupled issues, and often risk a simplification of the risk spectrum. Care should be taken when using quantitative probabilities and historical statistical data [54] that hide the uncertainty of risk assessment behind a veil of mathematical certainty. For the reasons explained above, classic probabilistic methods are not adapted to identify "risk scenarios in case of highly complex, dynamic, hybrid systems of hardware, software, and human components" [49] (p. 8), which are exactly the Industry 4.0 challenges identified in this paper. "Digital systems may be at intermediate fault modes before reaching a final failure state that will be revealed to human operators in the humane machine interface" [55] (p. 25), and lead to unplanned, unfamiliar, invisible and incomprehensible unexpected sequences of failures. It should be recognized that other more nuanced methods, that also use risk indices, exist and are meanwhile applied in the process and petrochemical industry; these nuanced methods were not referred to among our consulted sources, as explained in Section 2, materials and methods. Some examples that sometimes lead to richer interpretations of safety are: (i) the inherent safety index [56], which assesses the principles of inherent safety, being minimization, moderation, substitution and simplification; the (ii) safety weighted hazard index, which assesses a process unit of an industry by simultaneously evaluating hazards and hazard control measures [57]; or (iii) fuzzy indices that account for uncertainties in hazard and risk calculations [56].

3.2. Complexity-Thinking-Related Safety Paradigms and Concepts

This section is divided into several subsections that introduce the reader to paradigms, principles and concepts that either have complexity thinking as a central premise or are closely related to it.

3.2.1. From Linear Decomposition to Non-Linear Systems Causality

Complexity thinking marks a changing perspective on causality, moving from sequential models to systemic models [11], which is a change from linear thinking to non-linear thinking and from an individual to a holistic perspective. In the complexity-thinking paradigm, “failure is the result of the adaptations necessary to cope with the complexity of the real world, rather than a breakdown or malfunction” [58] (p. 167). Systems thinking is often applied inseparably from complexity thinking. Both schools of thought highlight different but intertwined aspects in socio-technical systems. Systems thinking marks the changing perspective from decomposition by analytical reduction to the analysis and design of the whole, as distinct from the components. It provides a means for studying emergent system safety properties [59]. This overcomes feedback loop problems, non-linear interactions between components and the fact that interactions among the sub-systems cannot always be considered separate from the behavior of the sub-systems themselves [59]. Note that these paraphrases, which we borrowed from Leveson in regards to systems thinking, are strongly related to complexity thinking, because the holistic approach of seeing the system as a purposeful whole and not as a collection of parts has an immediate effect on the complex interactions within this system. Several of the traditional methods from Table 1, like FMEA, fault and event tree analysis and root cause analysis have paradigms that are based on the decomposition of events and systems and differ from the systems-thinking perspective.

Several extracts from the 1980s start to cover different aspects of systems-thinking and complexity-thinking theories. Turner (1978) [60] was the first to write about fallible constructs in previous eras of safety thinking, followed by Perrow (1984) [61] with a critical and sometimes pessimistic view on developments of out-of-control technology, and Rasmussen who started a tradition of self-organized emergent systems (1980s and onwards) [62]. All of these authors still have a strong influence on contemporary safety science. Perrow wrote his book *Normal Accidents* in 1984 as a reaction to the fallibility of nuclear power plants after the Three Mile Island incident. Nuclear power plants were initially assessed to be ultra-reliable, but failed on a socio-technical level anyway.

3.2.2. Complexity Expressed as a Combination of Tight Couplings and Non-Linear Tractability

Although our scope is completely different, Perrow’s thoughts are still used today to deconstruct socio-technical systems with different degrees of complexity and coupling [14,16,63–65]. Complexity is expressed in terms of linear versus complex interactions and coupling discriminates against the possibility of either loosely or tightly coupled systems. Tight couplings mean that tasks or events are tightly connected or dependent [61]. A consequence of tightly coupled systems is that “there is no slack or buffer or give between two items” [61] (p. 90), and they are more difficult to control due to system-wide failure propagation [11]. Loose coupling means that tasks or events are independent and the failure of one component does not affect the failure of another component, nor does it cause the breakdown of the whole system.

Perrow presented a diagram, based on his breakdown of complexity and coupling, resulting in four quadrants. (See Figure 2 for a more recent revision [14] by Erik Hollnagel of Perrow’s original diagram.)

Different areas of industrial and commercial activities are positioned within one of these quadrants as a result of the combination of coupling and complexity characteristics. Perrow positioned manufacturing and assembly lines in the third quadrant, being defined as loosely coupled and resulting in linear interactions. Nuclear power plants (NPPs) can be found in the most critical, upper right position, and were the immediate motivation for Perrow’s book. With the introduction of Industry 4.0 characteristics both couplings and complexity increased substantially in many industrial processes. Hollnagel stated in 2008 that shifts towards more coupling and complexity for certain industries, and hence changes in quadrant positioning, are to be expected [14]. The technologies of Industry 4.0 act as a turbocharger for the further tightening of couplings. Factories have become “‘smart,’ i.e., they are highly efficient in resource, and they adapt very quickly to meet management goals and current industrial scenarios” [4] (p. 408). Physical systems are becoming integrated in networks: “the information technology part of Industry 4.0 consists of cyber-physical systems (CPS), cloud computing, and the internet of things (IoT)" [4] (p. 408), the latter being defined by Xu et al. [66] as machine–machine interactions without human intervention [4]. Whereas the traditional risks of robots in confined and protected spaces were relatively easy to identify and control, "more flexible and mobile cobots (a hybrid of the words 'collaborative' and 'robots' that refers to robotic devices that manipulate objects in collaboration with a human operator [67]) performing all sorts of tasks in close interaction with workers represent a much broader range of much less predictable risks" [13] (p. 408). Increased human–machine interaction "does not mean simply more human–machine control interfaces, but new ways of sharing tasks in order to complete complex operations more rapidly" [13] (p. 407). Although Industry 4.0-incentivized decentralization can meet the growing needs of highly customized products [4], it also adds non-linear interactions to work systems.

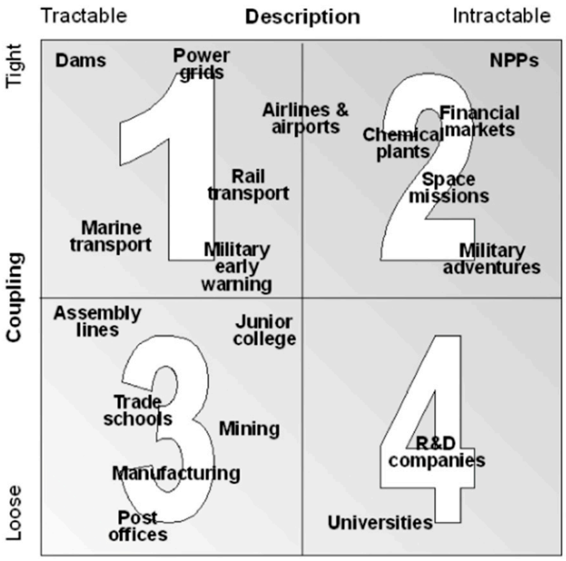

Figure 2. The coupling/tractability diagram by Hollnagel (2008), revised from Perrow's original (1984) (with permission from Erik Hollnagel).

3.2.3. Shift from Safety-I to Safety-II

The shift from simple to complex systems has recently been complemented by a shift from Safety-I to Safety-II thinking. Safety-I defines safety as a condition where the number of adverse outcomes (accidents/incidents/near misses) is as low as possible [58]. In order to improve safety, one 'needs' adverse outcomes to facilitate the learning process. This has been a well-proven strategy since the beginning of safety management. In particular, for 'safe' industries such as aviation, the approach is limited because of an absence of incidents and a subsequent absence of learning opportunities; this is called the paradox of safety [42]. Even in less-safe industries, failures represent just a small fraction of normal performance in which things often go right. Consequently, learning from failures only creates limited learning potential, nor does workplace safety simply equal the absence of work-related injury [45]. Therefore, Safety-II focuses on a system's ability to succeed under varying conditions, so that the number of intended and acceptable outcomes is as high as possible. Because safety in this paradigm is a system's ability, an organization and its operators cannot be safe or unsafe, but need to constantly navigate success under limited resources, goal conflicts and trade-off decisions. From this follows the principle of 'approximate adjustments' [68], which is the assumption that people continuously adjust what they do so that their actions match their conditions. This is in fact the theoretical reason why things more often go right rather than go wrong. The scope of analysis under Safety-II is therefore normal performance variability from everyday work, which has a very different etiology than attributing binary outcomes of safety, as found in Safety-I [58]. Instead, outcomes emerge from human performance variability, which is the source of successes as well as failures. This is called the principle of 'equivalence' [68]. While some adverse outcomes can be attributed to failures and malfunctions, others are best understood as the result of coupled performance variability. Therefore, Safety-II is not intended to replace Safety-I, but extends Safety-I [58,69].

3.2.4. Causation Model Concepts and Their Relation to Coupling and Tractability

In 2008, Hollnagel revisited Perrow's framework to meet the demands of increasingly complex socio-technical systems in relation to accident investigation methods. Hollnagel warns that whereas Perrow's notion of coupling is relatively straightforward, the notion of complexity must be used with care, since there is a difference between the complexity of an actual system and the complexity of its representation. Instead, Hollnagel proposes to replace complexity with tractability, meaning "how easy it is to manage or control the system" [14] (p. 7). "A system, or a process, is tractable if the principles of functioning are known, if descriptions are simple and with few details, and most importantly if the system does not change while it is being described. Conversely, a system or a process is intractable if the principles of functioning are only partly known or even unknown, if descriptions are elaborate with many details, and if the system may change before the description is completed" [14] (p. 7). Hollnagel extended Perrow's framework by positioning accident investigation methods into the quadrants of coupling and tractability, based on the ontology of the underlying causation model. See Figure 3 below.

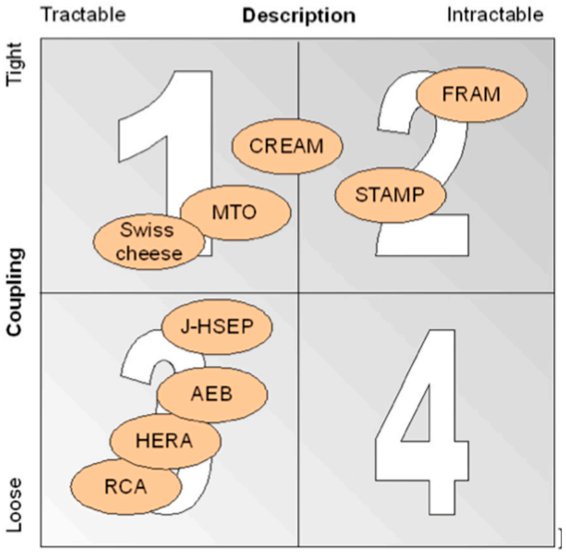

Figure 3. Characterization of accident methods (with permission from Erik Hollnagel).

Although Figure 3 depicts accident investigation methods, Hollnagel attributed the reasons for categorizing their coupling and tractability capabilities to the underlying causation models. Prospective safety analysis methods have the same causation models and we thus created a similar matrix with the methods from Table 1 (and the complexity methods that will be explained in a later section, see Table 2 in Section 3.2.6) by matching the underlying causation models to their prospective counterparts in safety analysis methods. These causation models can be retrieved from the paradigm label that is assigned for each method from Table 1.

In relation to the paradigms 'energy barrier thinking', 'linear causality', 'single cause philosophy' and 'human reliability assessment' (HRA) from Table 1, Hollnagel et al. defend that the following categories of methods affect the analysis of loosely coupled and tractable systems [14] (p. 11): "(i) methods that focus on the identification of failed barriers, (ii) methods that focus on human error, (iii) methods that focus on root causes in isolation, and (iv) methods that focus on root causes in combination". In accident analysis methods this includes accident evolution and barrier (AEB) analysis, a specific human error method (HERA), a multiple root cause analysis method (J-HSEP) and root cause analysis (RCA). In prospective safety analysis, this coincides with most of the method paradigms, because the majority is based on root cause analysis, human error assessment and energy-barrier thinking (see Table 1 for the labels and Table 4 for the resulting matrix). Although several traditional methods are not suited to analyze intractability issues, some of them are adapted to handle some types of tight couplings. FMEA is especially designed to analyze the reverberations from cascading failures, and HAZOP assesses the effects from deviating parameters on the net result of process safety. Although LOPA and Bowtie analysis are methods based on energy-barrier thinking, they take into account the couplings between several components and their layers of defense. Why Hollnagel assesses human error assessment methods to quadrant 1 deserves some additional justification in the section below.

3.2.5. Causation Models and the Human Contribution

Although there was a shift in focus from the failure of technical systems since the first causation models from the middle of the 20th century towards human–machine or socio-technical failures by the end of the 20th century, this largely resulted in a search for human failure as an oversimplified conception of human error [70]. The small impact of human reliability approaches on many industries is attributed to a failure to consider the challenges of systems development [71]. Human reliability assessment (HRA) has additionally received an increased critique for displaying benchmarking problems between methods with great variations in outcomes, poor methodological support, poor support for error prediction and analyst subjectivity [71,72]. In HRA, human cognition is assessed the same way that technical failure is assessed, by decomposition of systems and by assigning probabilities to sequential cause–effect events. Organizational, political, economic and environmental factors that influence human cognition are outside the consideration of HRA methods [46], and the contextual validity of such methods is generally low [71]. Human performance is always context-sensitive [70], which makes it either easily intractable or tightly coupled, or possibly both.

This has to some extent been recognized through the additional consideration of latent failures in accident investigations since the late 1980s. The Swiss cheese model has been instrumental in recognizing failures as the outcome of a "combination of active failures and latent conditions, rather than as the result of isolated events at the sharp end" [73] (p. 16), and hence as a counteraction to the oversimplified conception of human error. In this and other epidemiological models, dormant failures from design, maintenance and management can converge to a catastrophic event by combining active operational failures under specific system conditions [55]. Although in epidemiological models context-sensitive information becomes relevant for understanding human error, it also remains difficult to reach consensus on the contextual sources of latent failures [71]. Epidemiological models are also called complex linear models [74]. They are both graphically and conceptually an update of the Domino Model. The term 'complex' refers to the fact that multiple active and latent conditions must occur in combination for an accident to happen; the word 'linear' refers to the fact that the Swiss Cheese Model still depicts a successive sequence of events and does not yet explain emergent systems behavior. These two terms could explain why Hollnagel has attributed the Swiss Cheese model to the first quadrant, as it is able to reflect tight coupling interaction (complexity of interactions) while remaining a method to solve tractable failures (linearity of causes and defenses). This is why the 'epidemiological causation' paradigm has been categorized as being capable of analyzing tightly coupled, tractable issues.

Note that although the Swiss cheese model is used as a framework both for accident investigation and as a safety communication model, it lacks the appropriate form for making predictions [73], and further lacks the properties to be considered as an actual method. However, due to its recognizable representation, it remains a clear example of an epidemiological causation model that finds its counterparts in Table 1's LOPA and Bowtie safety analysis methods. They can both be regarded as complex linear methods. Note that both are not just models, but methods, since they are more than just graphical representations and come with a structured series of steps to follow. Although Bowtie is sometimes described as simply being a combination of fault tree and event tree analysis [15], the methods' vocabularies makes use of threats that can be both latent and actual threats, and are comparable to the Swiss cheese representation. Pasman [15] writes that by applying indicators, statistics, Bayesian network, and expertly selected weight factors, both human and organizational factors for failure probability scenarios can be determined by the representation of Bowties. The integration of human and organizational factors is a clear reference to epidemiological models with their tightly coupled capabilities. Nevertheless, as all these conditions are assumed to be foreseeable as the combination of events, they are assessed as tractable.

One remaining paradigm that has not yet been explained is a "rationalist, prescriptive and top-down belief in procedures" [75], which is only one interpretation of two possible procedural paradigms. In the first procedural paradigm, "rules are seen as static, comprehensive limits of freedom of choice, imposed on operators at the sharp end and violations are seen as negative behavior to be suppressed" [75] (p. 207). "The second procedural paradigm says that procedures are socially constructed, locally situated and bottom-up; it is rooted in sociology and work ecology. It envisions procedures as emerging from work experience and recognises that they are essentially incomplete and require translation and adaptation to any specific situation" [76] (p. 165). Methods such as the action error method, which identifies departures from specified job procedures, or deviation analysis, which distinguishes deviations from the planned and normal production process, reveal a belief in maintaining stable conditions of work by adhering to a fixed process or static set of procedures, hence a belief in the tractability of systems. This belief is rooted in 'Taylorism' [76], another paradigm that is strongly associated with the first procedural paradigm. Taylorism assumes that work prescriptions and task breakdowns can be efficiently controlled by supervisors and middle-management as the best way to assure operator reliability [75,77]. Taylorism pre-supposes linearity between procedures and hazards, and therefore assumes strict tractability. In Tayloristic methods, just as in the first procedural paradigm, tight-coupling is not assessed, but at best avoided by predictable and reliable performance.

3.2.6. The Identification of Complexity-Thinking Methods

In "complex nonlinear interactions, failures do not arise from the relationship of component failure modes and their causes, but 'emerge' from the relationships between these components during operational situations" [55] (p. 26). The evolving Industry 4.0 concerns, as discussed in this paper, need a safety analysis method that is capable of analyzing tightly coupled and intractable work system issues; in other words, a method situated in the second quadrant. In line with Hollnagel [14], we agreed to apply the notion of tractability instead of complexity in the diagram. We refrained, however, from positioning our objects of investigation on a sliding scale within the quadrants. Perrow [61] described the likelihood for a bias of his choices. We therefore think it is difficult to defend how one method is relatively positioned in terms of being more or less tractable or tightly coupled in comparison to other methods within a quadrant on a sliding scale, and therefore decided to use a table instead. The following steps therefore consist of identifying methods that are capable of examining the dynamics of complex interactions in a systems theory-based approach. Figure 3, from the study from Hollnagel et al. [14], provides a first opportunity for identifying prospective safety analysis models and methods, since the accident investigation model System-Theoretic Accident Model and Processes' (STAMP) and the Functional Resonance Analysis Method (FRAM) that can be retrieved from the complexity quadrant (tightly coupled and intractable) are both prospective as retrospective models and methods. STAMP is the denominator for the common accident causation model from (i) the prospective hazard analysis technique, called systems theoretic process analysis (STPA), and (ii) the accident and incident analysis technique, called Causal Analysis based on STAMP (CAST). STAMP as a model thereby forms the basis for both a retrospective and prospective method. In between quadrant 1 and 2, one can additionally find the Cognitive Reliability and Analysis Method (CREAM), which can be applied both retro- and prospectively, but this method, although considered by many as rather successful and still in use, has been abandoned by the inventor Hollnagel [15] for several reasons. In Hollnagel's own words [78]: "first, because it focuses on how actions can fail, rather than on the variability of performance"; secondly, "because it focuses on one part or 'component' of the system only, namely the human(s)". Another reason is that CREAM was misinterpreted to accept the concept of human error. FRAM, a method that originated from the same creator, has currently replaced several shortcomings of CREAM [78]. Several publications in search of systemic and complexity-thinking-inspired accident causation methods have confirmed FRAM and STAMP as the most cited and most suitable methods [21,24,79]. Underwood et al. [80] produced a review of systemic accident analysis methods and concluded that STAMP, FRAM and Accimap were the most used systemic accident analysis methods. Accimap has no prospective counterpart, but Rasmussen's hierarchical risk model on which Accimap is based, has also been integrated in some prospective systemic complexity-thinking methods like STAMP, as well as a particular use of FRAM with an abstraction/agency extension [81]. The integration of the hierarchical socio-technical deconstruction in STAMP and the above-mentioned extended FRAM display an unmistakable link to systems thinking. Some authors mentioned Event Analysis of Systemic Teamwork (EAST) [24,82] as an additional, although less-cited method that meets the requirements for an analysis method that is able to assess tightly coupled and intractable systems. Visualization tools that did not meet criteria to be counted as full methods, such as the ones described by de Vries [83], were excluded, but will be discussed as interesting approaches in Section 3.3.1. It is beyond the scope of this paper to fully explain the methods, but they are fully examined in the original sources: STAMP [16,84], FRAM [68] and EAST [82,85]. A short description of each can be found below.

Event Analysis of Systemic Teamwork (EAST)

EAST uses a three-step approach of data collection, data analysis and network representation. Input data can be derived from several sources including interviews, questionnaires, observational data, communication transcripts, critical decision method probes or combinations thereof [82,85,86]. After this, task, information/communication and social networks are analyzed in isolation and in combination with the help of network metrics. Each network is analyzed with its own particular existing method. Subsequently, an aggregation of two or three networks is analyzed for additional resultant and emergent interactions. The first studies with EAST used hierarchical task analysis, critical decision methods, coordination demand analysis, communications usage diagrams and operation sequence diagrams, whereas later versions developed network information directly from the raw data [82]. In a last step, the outcomes from the network metrics are supported by various integrated graphical representations.

System-Theoretic Accident Model and Processes (STAMP)

STAMP is based on how safety constraints and process models interact in hierarchical safety control structures. Together, these three essential concepts produce a diagram where the safety control structures are organized hierarchically. All these levels are believed to have influences on each other in the form of reciprocal communication channels, beginning with the societal level, followed by the levels of organization factors, company management, operational management and, finally, the actual operating process structures. Within these levels, one finds process models where a controller controls a process through control actions and receives responses through feedback or measurement outputs. Finally, every controller has an internal control model that directs these feedback and control loops [16]. Failures occur when the process model used by the controller does not match the process as a result of incorrect or unsafe control commands, an absence of required control actions, wrong timing of control actions or the control being stopped too soon or applied too long [16]. The results are safety control structures that determine potentially hazardous control actions [79] on the micro, meso or macro level. Data gathering can have all sorts of inputs and is not prescribed by the method.

Functional Resonance Analysis Method (FRAM)

A FRAM analysis follows a four-step approach performance [68]: (i) modeling the system; (ii) identifying the inevitable variability of work-as-done (see Section 3.2.8); (iii) aggregating the variability; and (iv) managing the variability. Each action performed by an agent (individual, group, equipment, artefact or organization) [81] within a work system is described by a function depicted by a hexagon. Each corner of the hexagon represents the six fundamental aspects (i.e., input, output, time, control, precondition and resource). These aspects can be linked to the aspects of other functions and produce a systematic network representation of functional interactions. The potential for variability between these functions is assessed by endogenous, exogenous and/or upstream–downstream coupling. The variability of the aspects can also be described by various phenotypes such as timing, precision, speed, force, and so on. The variability of the model and its emergent behavior, called functional resonance, is assessed by analysis of the upstream–downstream interactions. To manage variability, positive resonance will be amplified, while negative resonance must be dampened. This is achieved by inserting barriers, rearranging the order of functions, assigning roles to other agents, creating redundancies or reorganizing the work system according to a better understanding of its functional resonance.

Table 2, displayed below, is the counterpart of Table 1, this time for complexity-thinking-inspired methods; it displays the same structure and labels as Table 1. Table 2 is immediately followed by Table 3, which now merges the coupling and tractability capabilities from methods contained in both Tables 1 and 2.

Table 2. Complexity-thinking-inspired safety analysis methods (concept, paradigm and basis for structuring).

Method	Concept	Paradigm	Basis for Structuring	Coupling/Tractability
STAMP [16,21,24,79,80,83,84]	Creation of a model of the functional control structure for the system in question by identifying the system-level hazards, safety constraints and functional requirements.	Feedback control system	Most basic element in the model is a constraint, whereas basic structuring is the feedback control system	Tight coupling—intractable
FRAM [21,24,68,79,80,83]	Systemic analysis of complex process dependencies, based on the idea of resonance arising from the inherent variability of everyday performance.	Functional resonance	Dependencies among functions or tasks	Tight coupling—intractable
Event Analysis of Systemic Teamwork (EAST) [24,82,85]	A means of modeling distributed cognition in systems via three network models (i.e., task, social and information) and their combination.	Propositional network	Task, social and information network connections	Tight coupling—intractable

Table 3. Coupling and Tractability of the safety analysis methods matrix.

	Tractable	Intractable
	Bayesian Networks	EAST
	Layer Protection Analysis (LOPA)	FRAM
Tightly coupled	Bowtie Analysis	STAMP
	Hazard and Operability Studies (HAZOP)	Sneak Circuit Analysis
	Failure Mode Effect and Analysis (FMEA)	Monte Carlo Simulation [1]
Loosely coupled	Energy Analysis	
	Fault Tree Analysis	
	Event (Effect) Tree Analysis	
	Action Error Method	
	Job Safety Analysis	
	Deviation Analysis	
	Safety Function Analysis	
	Change Analysis	
	Root Cause Analysis	
	Human Reliability Assessment	
	Deterministic Probabilistic Risk Assessment	
	(e.g., FN curves—Risk Indices)	
	Databases	
	(e.g., Reaction Matrix—Consequence Analysis)	
	Cognitive Task Analysis	
	Audits	

[1] when used in conjunction with complexity inspired methods (e.g., Patriarca et al. (2017) use a Monte Carlo evolution of the Functional Resonance Analysis Method (FRAM) to assess performance variability in complex systems).

Note that quadrant 4 in Figure 3 and Table 3 are empty for historical reasons [14]. The development of accident causation went from loosely coupled, tractable systems at the beginning of industrialization, to more tightly coupled systems, which automatically reduced tractability. Research has found that increasing coupling also increases operational complexity [63]. The development of investigation methods simply matched the development of the accidents.

3.2.7. Resilience Engineering

Resilience Engineering (RE) proposes yet another view on safety that is in line with the intractability credo due to emergent system behavior. RE is closely related to Safety-II thinking, but it merged from another school of thought. Its many definitions and interpretations were reviewed by Patriarca et al. [87], with one frequently used definition being "the ability of the system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required performance under expected and unexpected conditions" (Robson, as cited in [87]) (p. 87). RE thinking includes a non-binary view on safety outcomes, including (i) buffering capacity, the degree to which a system can respond to disruptions before it breaks down; (ii) brittleness, the system's ability to restructure in response to external pressures; (iii) margin, how closely the system is operating relative to its safety or performance boundaries; and (iv) tolerance, how systems behave near these boundaries [88]. Feedback loops and controls in STAMP can represent more than just binary states, including buffers, margins and tolerances from which the brittleness of the system emerges. The functional representations from FRAM have a similar non-binary goal of representing performance variability, and the total of functional resonances makes up the brittleness of the system. EAST applies network metrics (e.g., in social systems) to semi-quantitively assess the 'centrality', 'closeness', 'betweenness' and 'eccentricity' of actor relationships [82], which can be translated into buffers, margins and redundant tolerances and their resulting behavior. Such non-binary, non-deterministic interpretations of safety-display capabilities seek to manage the challenges of intractability. Tight couplings can be assessed by the functional aspects of FRAM as well as the above-mentioned social network metrics, supplemented by reception and emissions between agents in EAST. STAMP represents the relationships at the different socio-technical levels to show reverberations and information transactions throughout all levels of the work system. Both Safety-II and RE show that non-deterministic definitions of safety need suitable methods to assess safety not caused by a coincidence of independent failures but instead by "a systematic migration of organizational behavior to the boundaries of safe behavior" [16] (p. 52), including how such boundary conditions are challenged and change and over systems' life cycles, or how adaptive shortfalls occur and can be counteracted by increasing resilience. Even the way systems fail is part of resilience assessment. Systems can either gracefully degrade or show brittleness, the property of displaying sudden dramatic failure "when events push the system up to and beyond its boundaries for handling changing disturbances and variations" [89] (p. 5). Note that RE uses a completely different definition of risk than the one implied by many traditional safety analysis methods that define risk as the combination of the likelihood of occurrence and the severity of injury, being also the definition which ISO 45001 [32] proposes for an occupational health and safety risk. The analysis of traditional methods stops at the failure, whereas our complexity-thinking methods are capable of analyzing systems reactions to stressors before, during and after failure events. The understanding of safety in RE is radically different and belongs to another paradigm.

The traditional methods from Table 1 are not appropriate to assess such intractable, non-deterministic degrees of safety. However, probabilistic sneak circuit analysis, a hardware–software reliability method, forms a mentionable exception. It displays "latent hardware, software or integrated conditions that may cause an unwanted event to occur or may inhibit a desired event and is not caused by component failure. These conditions are characterised by their random nature and ability to escape detection during the most rigorous of standardised system tests" [26] (p. 70). Note that this definition matches both RE's search for margins and brittleness, as Safety-II thinking's focus is on performance variability; as advocated in Safety-II and RE, it avoids unwanted events and promotes desired events. Sneak circuit analysis is a typical example of a method that only assesses one small link in a chain of methods, and jointly assesses human-centered coordination and its effects in a socio-technical work system.

3.2.8. Work-as-Imagined versus Work-as-Done

With the introduction of RE and Safety-II thinking, there has been an increased interest for work-as-done (WAD) analysis. WAD is defined as what actually happens in work systems. "It is messy and completely context dependent. That context is described by the interplay between all the components of a complex, non-linear system" [90] (p. 408) that consists of a wide variety of interactions between operators, equipment, procedures, processes and working environments. Work-as imagined (WAI) on the other hand is defined as how we or others imagine work should be performed [76]. This includes managers and regulators, who often design procedures without being involved in the actual processes. WAD analysis has become crucial in relation to Safety-II to accurately describe work systems without the usual distortion brought by normative bias. As previously explained, we disqualified semi-structured interviews, Delphi technique and multi-criteria decision analysis from Table 1 for merely being tools and not classified as methods. Note, however, that they are appropriate tools in conjunction with any method to gather representative WAD information. In FRAM, the principle of approximate adjustments inherently includes the requirement for a WAD analysis, as otherwise it will not be possible to make an accurate analysis of how people manage the challenges and trade-offs in everyday work. Sometimes FRAM models from a WAD and a WAI perspective of the same work system can be compared [91,92] to better align procedures with the challenges of the actual work [93]. In STAMP, the analysis can be performed even in the design phase of a system, before a system is operational and when WAD is not yet present. Even if EAST does not make WAD analysis explicit in its methodology, it assesses task and communications networks from operational work systems, which again are derived from running systems.

3.2.9. Complexity-Thinking Critique from the Narrative Literature Review

To conclude, and in line with a narrative literature review [94], we also looked for literature that was critical to complexity thinking, systems thinking, resilience engineering, Safety-II or work-as-done analysis. We have not found any literature that tried to disassemble these concepts and paradigms, but only found marginal remarks. Hovden et al. [21] remind us that one should not forget that many hazards and risks in OHS remain the effect of loosely coupled and tractable failures, and that linear traditional methods also remain essential instruments for the analysis of particular hazards and risks. Smith et al. responded to an editorial article that advocated Safety-II approaches in healthcare, stating that neither Safety-I nor Safety-II will develop its full potential when professionals lack the skills and training they need in order to perform their duties in relation to the modern-day patient safety agenda. The authors have expressed reservations towards whether a mere shift from Safety-I to Safety-II will be sufficient [95]. Such a critique in healthcare could easily be translated to manufacturing. Likewise, Salmon et al. [24] observe that a continuously increasing presence of emergence, resilience, performance variability, distributed cognition and complexity in modern catastrophes stretches the capabilities of our analysis methods, including the systems ergonomics and complexity-thinking methods described in this paper.

3.3. Breakdown of Challenges

This section represents the yellow analysis segment of Figure 1, with a further consideration of the Industry 4.0 risks. Table 4 represents a further conceptual breakdown from our initial set of four important Industry 4.0 challenges: interconnectivity, autonomy, automation in joint human–agent activity and a shift in supervisory control. All of these challenges are human-centered socio-technical issues in line with the scope of our study.

Table 4. Breakdown of challenges.

Concept	Breakdown of Challenges	Label
Interconnectivity [1]	Oversimplifications in the face of the complexities of joint systems	Joint cognitive system as a system of distributed cognition with emergent behavior
	Oversimplification of functional allocation problems	Human–machine opposition fallacy
	Disintegrative units of analysis that separate humans, machines and interfaces	Separation fallacy
	Oversimplification of different degrees of substitution between people and automation given different levels of autonomy and authority of machines	Substitution myth
Autonomy [2]	Transform practice and coordination across human and machine roles	Envisioned word problem
	Create new kinds of cognitive work for humans, often at the wrong times; every automation advance will be exploited to require operational efficiency	The law of stretched systems
	Create more threads to track; makes it harder for people to remain aware of and integrate all of the activities and changes around them	Coordination costs
	New knowledge and skill demands are imposed on humans and humans might no longer have a sufficient context to make decisions, because they have been left out of the loop	Transformation of knowledge and expertise
	Coordinate/synchronize joint activity; make machine a team player. Team play with people and other agents is critical to success	Principles of interdependence
	Resulting explosion of features, options and modes creates new demands, types of errors and paths toward failure	Transparency of complex systems
	Machines, humans and macrocognitive work systems are fallible; errors are therefore systemic; new problems are associated with human–machine coordination breakdowns; machines now obscure information necessary for human decision making.	Principles of complexity
Automation in Joint Human–Agent Activity [3]	To be a team player, an intelligent agent must fulfill an agreement (often tacit) to facilitate coordination, work toward shared goals and prevent breakdowns in team coordination	Shared knowledge, goals and intentions that are committed to goal alignment
	To be an effective team player, intelligent agents must be able to adequately model the other participants' intentions and actions vis-à-vis the joint activity's state and evolution	Adequate (shared) models
	Human–agent team members must be mutually predictable	Predictability;
	Agents must be directable	directability
	Agents must be able to make pertinent aspects of their status and intentions obvious to their teammates	Revealing status and intentions
	Agents must be able to observe and interpret pertinent signals of status and intentions	Interpretation of signals
	Agents must be able to engage in goal negotiation	Goal negotiation
	Support technologies for planning and autonomy must enable a collaborative approach	Collaboration
	All team members must help control the costs of coordinated activity (note: the authors mean cost as in effort, not financial cost)	Coordination cost control

Table 4. *Cont.*

Concept	Breakdown of Challenges	Label
Shift in Supervisory Control [4]	Supervisor must have real as well as titular authority	Elimination of responsibility—authority double binds
	Supervisor must be able to redirect a lower-order machine cognitive system when the machine's problem solving breaks down	Operator has the authority to abort the operation; strategies for management of system boundaries
	Need for a common or shared representation of the state of the world and of the state of the problem-solving process	Adequate (shared) models

[1] Hollnagel and Woods (2006) [65]; [2] Bradshaw, Hoffman, Johnson and Woods (2013) [91]; [3] Klein, Woods, Bradshaw and Feltovich (2004) [94]; [4] Woods and Roth (1988) [90].

Actions of humans, machines and systems that jointly take part in socio-technical systems might be separated in time and space—something that is certainly true for Industry 4.0, where actions and cognition can be spread not only across actors, but even across business units. Therefore, approaches to assess individual risks will be limited and there is a need for systems-thinking approaches from which the theoretical underpinnings can be found (discussed in Section 3.2).

3.3.1. Interconnectivity

In relation to the challenges in safety-critical socio-technical systems in Industry 4.0, autonomous thinking and increased interconnectivity introduce the principle of distributed cognition as an omnipresent element in manufacturing and assembly environments, with an additional degree of complexity for human–machine cooperation. "Distributed cognition is the shared awareness of goals, plans, and details that no single individual grasps" [96] (p. 726), but which are distributed across the human and technical actors of the socio-technical system to jointly achieve a single goal. For collaborative robots (a typical example of a technology introduced by Industry 4.0), distributed cognition is even supplemented by the distribution of tasks. Hollnagel et al. have described analytical deceptions where systems are based on an "allocation of tasks between people and machines, which assumes decomposability of work into independent parts or tasks" [97] (p. 143). Oversimplification of functional allocation problems result in 'human–machine oppositions', which become subsequently translated into equivalent oppositions in safety analysis methods. This is defined as the 'separation fallacy' [97], where disintegrative units of analysis separate humans, machines and interfaces. "Ultimately the focus must be the design and the performance of the human–machine problem-solving ensemble—how to 'couple' human intelligence and machine power in a single integrated system that maximises overall performance" [98] (p. 422). Such an integrated approach can be found in joint cognitive systems (JCS) methods where the human–machine problem-solving ensemble becomes the unit of analysis of the investigation. Systemic models like STAMP, FRAM and EAST avoid the fallacy of human–machine opposition and are able to assess the joint system as the unit of analysis. STAMP ultimately assesses how the aggregation of many control structures result in the joint system behavior. FRAM uses a purely functional approach whereby each node represents a function regardless of whether its origin is anthropogenic or technological [68,91], and thereby avoids the use of mental constructs to assign cognitive labels in human-error-centered methods that are prone to oversimplification. Adriaensen et al. [91] used a JCS approach in FRAM to analyze aircraft cockpit systems and artefacts as agents on their own in order to understand the propagation of tasks between humans, systems and artefacts in an aviation scenario. EAST, STAMP and FRAM have replaced the notion of human error through a focus on the identification of human–task mismatches. EAST uses task, social and communication/information network models, not only in isolation but also by their joint effects [82], to display system behavior. "Cognition is therefore achieved through coordination between system units" [86] (p. 26), and ownership of information resides at the system level, not the individual level.

Such systemic models that are able to assess risk in distributed systems do not belong to the inventory of OHS environments [16,21]. Because of their fragmented units of analysis, traditional methods are prone to what some authors call 'the substitution myth' [99–101], where an oversimplification of different degrees of substitution between people and automation are allocated different levels of autonomy and authority as a static misconception of a dynamic reality. de Vries et al. [83] have dedicated a paper to the visualization of JCS and socio-technical systems models in prospective safety assessment and design. Besides STAMP, FRAM and activity theory (a simpler and earlier variant of EAST), de Vries et al. have listed a set of additional tools to overcome the different unit-of-analysis fallacies in Table 4.

3.3.2. Autonomy and Automation in Joint Human–Agent Activity

The challenges of autonomy and automation in joint human–agent activity are merged in a single section. 'Adequate shared models and knowledge' (See Table 4) [102] between human and machines are defined in STAMP by the analytical requirement of the internal control model of the controller, regardless of whether this controller is human or technological, to direct feedback and control loops [16]. These loops are defined in Table 4 as 'directability' (See Table 4) of machines and systems [102]. EAST and FRAM will display this as the result of functional exchange difficulties between agents or functions of the model. From a JCS perspective, automation, a key concept in Industry 4.0, acts as an additional agent with humanlike characteristics, which can be recognized by social metaphors applied to automation in the safety literature, including how to make automation 'team players' (See Table 4) [102] or how to certify the right amount of trust in automation [103,104]. Predictive judgement of machine competence is often the subject of mis-calibrated trust, leading to either excessive trust or mistrust [98]. One study examined the consequences when navigation robots apologized for their mistakes in trust repair [105], with human operators as one element, showing the intricacies of 'predictability' and 'team play' between humans and machines. The robot apology example also shows the significance 'to reveal status and intentions' (See Table 4) [102] for both machine and human.

Automation surprises as identified in the manufacturing industry [106] have a longer research legacy in healthcare and aviation. In some noticeable examples of these domains [107,108], automation surprises have caused or contributed to tragic catastrophes [109–111]. Automation interaction challenges in manufacturing list the same set of problems as the ones described in aviation accident reports. Mode transition surprises whereby "operators may become unaware of changes in the operating mode performed by automation" [106] (p. 455) are described in manufacturing automation. The same authors describe automation-induced errors, whereby more automation can induce new, unexpected forms of human performance or inappropriate distrust in automation from manufacturing environments. Just as in aviation, healthcare and operation of nuclear plants, manufacturing automation induces 'out-of-the-loop' conditions [106], which have been described as situations in which operators can encounter difficulties achieving a complete picture of the automation processes and comprehending how this may impair the detection of automation failures and the ability to regain manual control. In Table 4 this is defined as 'the transformation of knowledge and expertise' (See Table 4) [99], where new knowledge and skill demands might deprive human operators of sufficient context for making appropriate decisions and as 'the envisioned world problem' (See Table 4), which states that the introduction of autonomous technology transforms practice and coordination across human and machine roles Accompanying impaired detection from technology changes or transient 'out-of-the-loop' conditions can be translated directly in STAMP's reciprocal feedback and control loop mismatches, whereas FRAM identifies the absence of a precondition or control aspect to properly execute a downstream function and identify negative functional resonance. EAST can identify a mismatch between the task and information/communication network. The models are also capable of identifying gaps in transformed practices and coordination between roles for humans and machines and of identifying requirements for new knowledge and expertise at higher levels in hierarchical socio-technical systems. The absence of a precondition or input (FRAM), lack a control (STAMP) or flawed network relations (EAST) can also identify organizational aspects like suboptimal induction training or inconsistent operating manuals to accompany the introduction of new technologies.

The interconnectivity principle in Industry 4.0 induces interactive complexity, whereas other increasing forms of complexity [16] like dynamic complexity due to changes over time, or nonlinear complexity, where cause and effect show no obvious relation, should be assessed with methods that can assess intractable and tightly coupled issues. Methodologically, these coincide with 'the principles of interdependence' (See Table 4)and 'the principles of complexity' (See Table 4) [99] from Table 4, while simultaneously trying to achieve 'transparency of complexity' (See Table 4). Note that a descriptive WAD is paramount to resolving such issues, and that any WAI approach that expects operators to adapt to opaque systems will do little to alleviate the problem.

Woods reminds us that "[i]ncreasingly autonomous things such as road or airborne vehicles are not 'things' at all but instead are complex networks of multiple algorithms, control loops, sensors, and human roles that interact over different time scales and changing conditions" [112] (p. 132). Autonomous transport vessels can be found in some Industry 4.0 environments, but the networks of algorithms, loops and sensors are a basic premise of many other autonomous 'things' in the emerging Industry 4.0 work systems. Such interconnected networks create a potential for what researchers call 'strong silent automation', which is automation that fails to communicate "signals that allow operators to predict, control, understand, and anticipate what the machine is or will be doing" [99] (p. 2). Automation surprises, mode confusion, mis-calibrated mental models and literal-minded machines (defined as machines that correctly act upon an internal model that is not necessarily aligned with the actual model of the world [97]) have brought about a shift in the causes of accidents. Therefore, it is important for both humans and machines to commit to 'shared knowledge, shared models, goals, and intentions—committed to goal alignment' (See Table 4), and provide a clear 'interpretation of signals' (See Table 4) regarding status and intentions [102]. FRAM and EAST, in which functional resonance or network metrics provide deeper insights into the effects of interconnectivity, are well suited to apprehend network-like systems due to the correspondence between reality and method representations. EAST has explicitly been defined as suitable for the analysis of work systems that possess a common goal and coordination across different agents that are geographically dispersed, and is supported by numerous systems, procedures and technology [85].

All the putative benefits from automation designers often deliver promises of offloading both requirements for attention and a reduction in actual work. In reality, operators encounter novel cognitive demands, often at busy or critical times, and have a requirement to actively track multiple activities and changes [99]. This is what Wiener [113] defined as 'clumsy automation', a phenomenon that occurs when the benefits of technology occur during a low workload and the burdens imposed by it arise during periods of high workload or during the most critical phases of work, which is labeled in Table 4 as 'the law of stretched systems' (See Table 4) [99]. New technological capabilities lead surprisingly often to mis-engineered clumsy automation [114], and thereby paradoxically add complexity to socio-technical work practices. A WAD analysis can only be applied to a running system. Even if a WAD perspective (which is often associated with RE and Safety-II) is well suited to provide descriptive models and identify the effects of unintended design surprises, the guiding principle in the design of systems before they are operational requires a systems and complexity-thinking approach that makes the intractable tangible [104]. The many documented issues with 'strong silent automation' [99,107] and 'clumsy automation' [114–116] make clear how automation is a topic that is inherently intertwined with supervisory control.

3.3.3. Supervisory Control

In the middle of the 20th century, Paul Fitts [117] published what is now still called the 'Fitts list', producing 11 statements and thereby creating the first attempt of functional allocation. This principle is also often cited as the MABA MABA principle, which is the abbreviation of discriminating between strengths of men and machine: 'men are better at' and 'machines are better at'. In the MABA MABA principle, "human and machine are construed as actuating and information processing systems with different capabilities, on the basis of which it is possible to determine what should be automated and what not" [100] (p. 2). Although both the original Fitts list and later MABA MABA updates have been defended in the literature, they have also caused great debate (see [100,101]). There is currently no consensus that a static dichotomy between a functional allocation of man and machine is feasible. One of the conflicts of function allocation is that it results in ironies of automation, implying that automation transforms work practices in ways unanticipated by designers [100,111,118]. One of these ironies is that, despite advanced automation and autonomous systems, humans retain the ultimate responsibility [119] in the form of a last fail-safe mechanism: "Currently, the burden of decision making under uncertainty is placed solely on human operators" [120] (p. 6). This is why Table 4 refers to 'the elimination of responsibility—authority double binds' that "occur when a party has responsibility in that others may impose sanctions on that party following outcomes, yet that party no longer has sufficient authority to influence or control the processes that lead to outcomes" [97] (p. 153). Equally, systems must be designed so that the supervisor is able to redirect the systems cognition into lower-order cognitive tasks when the machine's problem solving breaks down [98]. Operators ought to swap between monitoring systems and take over when automation fails, which produces a "need for a common or shared representation of the state of the world and of the state of the problem-solving process" (p. 423). The action of monitoring, including the paradox of monitoring non-events [121], is anything but a passive task, and yet can be the predecessor to an escalation of reverting to manual control—something that was avoided in the first place because of the supposed supremacy of automation and is labeled one of the ironies of automation [118]. Supervisory control is therefore not simply a task where a human passively monitors a system, but an action in which the operator monitors, uses system knowledge, anticipates and, if necessary, acts. Therefore, the promise of automation substitution whereby "[a]utomation can replace human work without any larger impact on the system in which that action or task occurs, except on some measures of its output" [111] (p. 90) has been called a myth instead of a promise by several authors [99–101]. One particular study applied EAST to compare the role of the driver from traditional cars with automated driving systems and found that "there are considerably more cognitive tasks in the passive monitoring role than in the active driving role" [122] (p. 2).

4. Discussion

Traditional OHS safety analysis methods lack the ability to analyze socio-technical issues collectively, as discussed in Section 3.1. The closest attempt that has been made involved a number of methods that assessed human reliability, with a consideration of human operators as being part of systems that would work perfectly fine without their intervention [123]. However, this mechanical cause–effect approach towards the assessment of human performance, in line with the reliability of components, has been abandoned by several researchers [55,70]. It has been contradicted by the Safety-II paradigm, where human performance is now assessed in terms of approximate adjustments to mismatches in systems ergonomics while managing dynamic trade-offs and goal conflicts with finite resources [77]. The answer to the research question, "Can complexity thinking contribute to progress in occupational safety in Industry 4.0?" suggests a positive conclusion for several reasons. First of all, traditional methods fall short in identifying and assessing emergent system properties, whereas complexity methods respond to the absence of approaches that can assess tightly coupled intractable issues. In complexity-thinking methods, humans, machines and interfaces need not to be decomposed in disintegrative units of analysis. Contrarily, the methods are capable of taking the joint problem-solving ensemble as the unit of analysis. Consider, however, that this remains a researcher's choice, and the substitution and opposition fallacies from Table 4 are not automatically solved by applying complexity methods.

An important contribution from complexity thinking also comes from abandoning the identification of hazards and defenses as the only necessary measure to make fully protective systems. Instead, we suggest a shift to be able to better understand and explain the unique conditions of every work system. This is another reason why our initial choice for the phrase 'safety analysis methods' was adequate, because it comprises more than just hazard identification and risk assessment. FRAM, STAMP and EAST have been instrumental in better explanations of work systems, and they are able to better manage and mitigate risks from individual weaknesses or system brittleness. In FRAM, this would be the amplification of positive resonance and the dampening of negative resonance; in STAMP, it would be the reconsideration of control constraints and feedback loops; and in EAST, it would be to modify the causes of flawed network interactions. How to achieve these goals remains an individual choice for the user of a methodology. It would, however, be an oversimplification to think that complexity can be managed with predefined fixes.

4.1. Current Use of Complexity-Thinking-Inspired Methods for Industry 4.0

This raises the question of whether these methods are already in use in the context of Industry 4.0. When one looks to the previous uses of the EAST, STAMP and FRAM methods in industrial environments and autonomous and automated systems, relatively few publications have shown immediate connection to our topic. The combinations of 'Industry 4.0', or the original German term 'Industrie 4.0', with the names of the complexity methods provided zero returns in the Scopus database. When performing a broader search regarding EAST we mainly retrieved papers that concerned distributed cognition of crew work in aviation, maritime and military operations, and we retrieved some studies applied to road traffic interactions. The only paper that applied EAST with an appropriate correction to our topic concerned the changing role of the driver within automated driving systems [122]. This at least establishes the capability for assessing one of our main Industry 4.0 challenges: shifting supervisory control. STAMP was slightly more affiliated with the scope of our topic in previous studies. We found a study about STAMP/STPA hierarchical structures in risk analysis of a complex multi-robot mobile system, and an STAMP/STPA for software-intensive systems [124]. Most relevant was a STAMP case study regarding safety and security of a cyber-physical system [125], an actual Industry 4.0 application, and a clear example of an assessment of one additional challenge from our study: interconnectivity. Note that cybersecurity is another emerging challenge in Industry 4.0. [5,13], and although mainly a security issue, it impacts safety. Although FRAM is mainly used in maritime, aviation and health care operations [126], the method has been applied to analyses in industrial settings [127,128]. We identified an exploratory exercise that applied FRAM to manage OHS in complex and unpredictable manufacturing systems, although in a low automation environment [129]. FRAM was also used to refine operating guidelines in the manufacturing process of aeroengine blade forging [130]. One study compared fault trees, Bayesian networks and FRAM [46] applied to an industrial accident, and identified some advantages of FRAM, such as the ability to assess context-sensitive information, which the two other methods proved to be incapable of achieving. However, this FRAM study was a retrospective analysis, whilst our scope focused on prospective risks; in any case, the literature identifies that FRAM can be equally applied to retro- and prospective analyses. Even if our results are not exhaustive, we can learn from the relatively low content of relevant scientific papers that it is not only Industry 4.0 that needs to embrace complexity methods, but that complexity methods will need to embrace Industry 4.0.

Recently a study investigated the utility of integrating principles from the STAMP method with the EAST framework [131] to assess safety management in the design and operation of a railway level crossing. In relation to building FRAM models, progress can also be expected from the integration of Industry 4.0-like solutions. Wearables could be introduced to track operators' normal performances in everyday work, for example, as the basis for FRAM models concerned with variability in everyday activities. Likewise, software-based tracking of information exchanges could yield better understandings of the interactions between components of work systems. Artificial intelligence-based algorithms are also capable of supporting systematic and dynamic operator decision making in both normal and abnormal practices [132]. Although this paper has mainly discussed the underestimated threats from autonomy and automation in Industry 4.0, we should not forget that many Industry 4.0 technologies can be deployed for the benefit of safety as well.

4.2. Practicability of Complexity-Thinking-Inspired Methods for Industry 4.0

One needs to recognize that EAST, STAMP and FRAM methods do not have the benefit of a longer tradition in industry, and users first need to familiarize themselves with each methodology; an obstacle that once was also true for methods like HAZOP and FMEA. Complexity-thinking methods are not intuitive methods that can be used without studying their methodologies first. In terms of practicability, STAMP/STPA [133] and FRAM [126,134,135] offer open-source software tools to build models. Some EAST studies report the use of existing software previously designed for other network metrics analysis [82]. Even with the support of software, analysis using one of these three methods is very time-consuming. Some additions to the traditional FRAM model have included the use of the abstraction/agency framework [81], a framework based on Rasmussen's hierarchical risk model that deconstructs (rather than decomposes) socio-technical systems. Note that Rasmussen's socio-technical framework is already integrated in the STAMP framework. The myFRAM software also offers another addition to the traditional FRAM by offering a matrix representation [136] of dependencies. This matrix analysis makes it easier to systematically assess positive and negative functional resonances in the model and where they occur. A last addition to FRAM is a Monte Carlo simulation [47,48] to apply probabilistic distributions to scenarios and parameters of WAD models, which turns FRAM into a semi-quantitative approach/application and provides one way to deal with uncertainty. ISO 31010 lists Monte Carlo simulation as a risk management technique. Although it is mainly used to assess assumptions about frequency and the severity of incidents from classical risks, it can also be applied to estimate uncertainty of complex dependencies when used in conjunction with FRAM. A proposition to transform STAMP into a semi-quantitative engineering approach has been recommended by Karanikas et al. [137]. We suggest that consideration be given to the fact that EAST is inherently a semi-quantitative approach, as it applies network analysis metrics to assess a socio-technical system. Finally, it is worth mentioning that FRAM can be used in combination with the multi-criteria decision method AHP (see Section 3.1) to allow the simultaneous participation of multiple experts during the different steps of a risk analysis. The use of AHP to assign different phenotypes in FRAM reduces subjectivity [48] when determining performance variability. This approach has been used in FRAM applied to sustainable construction [138].

It is also valuable to consider that there are advances in traditional safety analysis methods by using dynamic and semi-automated hazard identification methods [15], although these techniques will not overcome their causation model limitations. Progress has been made by computer-coded stored 'flaw and failure scenarios', defined by traditional HAZOP and FMEA methods to question what can make a condition or item fail [15], rather than just determining a failure; this closes the gap with more complex methods. Unfortunately, complexity-thinking-based methods like EAST, FRAM and STAMP are very resource- and time-consuming, and it is not realistic to expect such methods to be used in the identification of an exhaustive list of industrial risks. Rather, the safety community needs to think about ways to prioritize which complex risks deserve a deeper analysis and how particular strengths and weaknesses of models can deliver the most efficient and resource-friendly solutions. Hans-Ringdahl's initial criteria for traditional safety analysis methods included that the application of the methods under consideration should be possible with reasonable resources and should be easy to understand for the user. In addition to being more resource-consuming than traditional methods, EAST, STAMP and FRAM will also require a more elaborate investment in the informants involved in an analysis. This is also true for the investigators in charge, who must become familiar with these methods. We suggest transforming Hans-Ringdahl's initial criteria into a criterion of proportionality with the complexity of the hazards and risks under consideration.

5. Conclusions

From the review presented in this paper, it has become clear that current safety analysis methods in OHS have some limitations and were mostly designed for risks established from another era. At the same time, complexity methods have been described as very time and resource intensive. Understanding and managing risks will therefore increasingly become a skill of switching between micro, meso and macro understandings of systems, which justifies the use of hybrid methods at different levels of abstraction from systems and their subsystems. There is also a need to unravel the current overlaps between hazard identification, risk assessment, injury models, management techniques and risk mitigation applied to guidelines for integrated approaches [31,39] as a countermeasure to cherry-picking methods. All in all, there is no one-size-fits-all solution, and risk management should not be an exercise of an isolated method or solution, but an integrated attempt to improve systems. The process industry with its increasingly tightly coupled functioning and network-like behavior will benefit from the complexity-thinking methods that today are still predominantly applied within an academic context. There will be a need to assess complex safety challenges, either to make the interactions in systems transparent or to examine specific safety issues of concern that can be managed with reasonable resources. Preferably, these methods should already find their application in the design phase.

In conclusion, we agree with the concerns from the literature that provoked the thinking and rationale found in this paper and that ultimately questioned whether OHS consequences of Industry 4.0 are being appropriately evaluated; we considered and provided a way forward via the introduction of complexity-thinking and systemic safety analysis methods. In addition, we also expect that some of the technical progress in Industry 4.0 can also be used for a better understanding of performance variability in work systems, and we call for further research that will support such developments. A magic bullet method does not exist, which is inherently related to the ever-varying conditions of work systems. To quote Vincent et al. [139] (p. 5): "Safety is a constantly moving target."

Author Contributions: Conceptualization, A.A., L.P. and W.D.; Methodology, A.A., L.P. and W.D.; Validation, A.A.; External validation support, P.R. (see acknowledgement); Writing—Original draft preparation, A.A.; Writing—Review and editing, L.P. and W.D.; Supervision, L.P. and W.D., Visualization, A.A.

Funding: This research received no internal or external funding.

Acknowledgments: We acknowledge Erik Hollnagel for his kind permission to use the coupling-tractability figures from his previous work. We also acknowledge Raphaël Weuts, our fellow researcher, for providing his academic expertise on Industry 4.0 and his support in drafting the methodology and revision of the manuscript. We are very thankful for the validation of the data and revision of the manuscript by Riccardo Patriarca from the Sapienza University of Rome, Italy, and we would like to thank Laura Maguire from Ohio State University, US for her efforts in the additional identification of relevant papers regarding theoretical concerns in relation to NER in Industry 4.0. We acknowledge Jozef Van Ballaer for his proofreading of the manuscript and for providing comments on the relevance of our findings from a practical Health & Safety manager perspective. Finally, we are very grateful to Judith Anderson and Frank Brehany for their corrections of the English language.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lasi, H.; Fettke, P.; Kemper, H.-G.; Feld, T.; Hoffmann, M. Industry 4.0. *Bus. Inf. Syst. Eng.* **2014**, *6*, 239–242. [CrossRef]
2. Kagermann, H.; Wahlster, W.; Helbig, J. *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Final Report of the Industrie4.0 Working Group*; Acatech: Munich, Germany, 2013.
3. Lee, J.; Davari, H.; Singh, J.; Pandhare, V. Industrial Artificial Intelligence for Industry 4.0-Based Manufacturing Systems. *Manuf. Lett.* **2018**, *18*, 20–23. [CrossRef]
4. Kamble, S.S.; Gunasekaran, A.; Gawankar, S.A. Sustainable Industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. *Process Saf. Environ. Prot.* **2018**, *117*, 408–425. [CrossRef]

5. Hermann, M.; Pentek, T.; Otto, B. Design Principles for Industrie 4.0 Scenarios. In Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA, 5–8 January 2016; pp. 3928–3937.
6. Lu, Y. Industry 4.0: A survey on technologies, applications and open research issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10. [CrossRef]
7. Miranda, J.; Ponce, P.; Molina, A.; Wright, P. Sensing, smart and sustainable technologies for Agri-Food 4.0. *Comput. Ind.* **2019**, *108*, 21–36. [CrossRef]
8. Winkelhaus, S.; Grosse, E.H. Logistics 4.0: A systematic review towards a new logistics system. *Int. J. Prod. Res.* **2019**, 1–26. [CrossRef]
9. Harms-Ringdahl, L. *Safety Analysis, Principles and Practice in Occupational Safety*, 2nd ed.; Taylor and Francis: London, UK; New York, NY, USA, 2005.
10. Waterson, P.; Robertson, M.M.; Cooke, N.J.; Militello, L.; Roth, E.; Stanton, N.A. Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety. *Ergonomics* **2015**, *58*, 565–599. [CrossRef]
11. Hollnagel, E. The Changing Nature Of Risks. *Ergon. Aust. J.* **2008**, *22*, 33–46.
12. Fernández, F.B.; Pérez, M.Á.S. Analysis and Modeling of New and Emerging Occupational Risks in the Context of Advanced Manufacturing Processes. *Procedia Eng.* **2015**, *100*, 1150–1159. [CrossRef]
13. Badri, A.; Boudreau-Trudel, B.; Souissi, A.S. Occupational health and safety in the industry 4.0 era: A cause for major concern? *Saf. Sci.* **2018**, *109*, 403–411. [CrossRef]
14. Hollnagel, E.; Speziali, J. *Study on Developments in Accident Investigation Methods: A Survey of the State-of-the-Art*; SKI Report 2008:50 (Swedish Nuclear Power Inspectorate); CNRS: Paris, France, 2008; ISSN 1104-1374.
15. Pasman, H.J.; Rogers, W.J.; Mannan, M.S. How can we improve process hazard identification? What can accident investigation methods contribute and what other recent developments? A brief historical survey and a sketch of how to advance. *J. Loss Prev. Process Ind.* **2018**, *55*, 80–106. [CrossRef]
16. Leveson, N.G. *Engineering a Safer World: Systems Thinking Applied to Safety*; MIT Press: Cambridge, MA, USA; London, UK, 2011.
17. Fabiano, B.; Reverberi, A.P.; Varbanov, P.S. Safety opportunities for the synthesis of metal nanoparticles and short-cut approach to workplace risk evaluation. *J. Clean. Prod.* **2019**, *209*, 297–308. [CrossRef]
18. Flaspöler, E.; Reinert, D.; Brun, E. *Expert Forecast on Emerging Physical Risks Related to Occupational Safety and Health*; European Agency for Safety and Health at Work: Bilbao, Spain, 2009; pp. 176–198.
19. Brocal, F.; Sebastián, M.A. Identification and Analysis of Advanced Manufacturing Processes Susceptible of Generating New and Emerging Occupational Risks. *Procedia Eng.* **2015**, *132*, 887–894. [CrossRef]
20. Onnasch, L.; Wickens, C.D.; Li, H.; Manzey, D. Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Hum. Factors* **2014**, *56*, 476–488. [CrossRef]
21. Hovden, J.; Albrechtsen, E.; Herrera, I.A. Is there a need for new theories, models and approaches to occupational accident prevention? *Saf. Sci.* **2010**, *48*, 950–956. [CrossRef]
22. Hulme, A.; Stanton, N.A.; Walker, G.H.; Waterson, P.; Salmon, P.M. What do applications of systems thinking accident analysis methods tell us about accident causation? A systematic review of applications between 1990 and 2018. *Saf. Sci.* **2019**, *117*, 164–183. [CrossRef]
23. Dekker, S.; Cilliers, P.; Hofmeyr, J.-H. The complexity of failure: Implications of complexity theory for safety investigations. *Saf. Sci.* **2011**, *49*, 939–945. [CrossRef]
24. Salmon, P.M.; Walker, G.H.; GJ, M.R.; Goode, N.; Stanton, N.A. Fitting methods to paradigms: Are ergonomics methods fit for systems thinking? *Ergonomics* **2017**, *60*, 194–205. [CrossRef]
25. Yousefi, A.; Rodriguez Hernandez, M.; Lopez Peña, V. Systemic accident analysis models: A comparison study between AcciMap, FRAM, and STAMP. *Process Saf. Prog.* **2018**, *38*. [CrossRef]
26. ISO 31010. *Risk Management-Risk Assessment Techniques*; International Organization for Standardization: Geneva, Switzerland, 2009.
27. Lees, F.P. *Loss Prevention in the Process Industries: Hazard Identification, Assessment and Control*; Butterworths: Oxford, UK, 1980.
28. Johnson, W.G.; Council, N.S. *MORT Safety Assurance Systems*; Marcel Dekker, Incorporated: New York, NY, USA, 1980.
29. Organisation, I.L. *Major Hazard Control: A Practical Manual: An ILO Contribution to the International Programme on Chemical Safety of UNEP, ILO, WHO (IPCS)*; International Labour Office: Geneva, Switzerland, 1988.

30. Bahr, N.J. *System Safety Engineering and Risk Assessment: A Practical Approach*; Taylor and Francis: Washington, DC, USA, 1997.
31. Khanzode, V.V.; Maiti, J.; Ray, P.K. Occupational injury and accident research: A comprehensive review. *Saf. Sci.* **2012**, *50*, 1355–1367. [CrossRef]
32. ISO 45001. *Occupational Health and Safety Management Systems Requirements with Guidance for Use*; International Organization for Standardization: Geneva, Switzerland, 2018.
33. Reis, M.; Gins, G. Industrial Process Monitoring in the Big Data/Industry 4.0 Era: From Detection, to Diagnosis, to Prognosis. *Processes* **2017**, *5*, 35. [CrossRef]
34. De Felice, F.; Petrillo, A.; Di Salvo, B.; Zomparelli, F. Prioritising the Safety Management Elements Through AAHP Model and Key Performance Indicators. In Proceedings of the Conference on Modeling and Applied Simulation, Larnaca, Cyprus, 26–28 September 2016.
35. Chan, A.H.S.; Kwok, W.Y.; Duffy, V.G. Using AHP for determining priority in a safety management system. *Ind. Manag. Data Syst.* **2004**, *104*, 430–445. [CrossRef]
36. Antonsen, S. Safety Culture Assessment: A Mission Impossible? *J. Contingencies Crisis Manag.* **2009**, *17*, 242–254. [CrossRef]
37. Le Coze, J.C. How safety culture can make us think. *Saf. Sci.* **2019**, *118*, 221–229. [CrossRef]
38. Dekker, S.; Nyce, J.M. There is safety in power, or power in safety. *Saf. Sci.* **2014**, *67*, 44–49. [CrossRef]
39. Pasman, H.J.; Rogers, W.J.; Mannan, M.S. Risk assessment: What is it worth? Shall we just do away with it, or can it do a better job? *Saf. Sci.* **2017**, *99*, 140–155. [CrossRef]
40. Lundberg, J.; Rollenhagen, C.; Hollnagel, E. What-You-Look-For-Is-What-You-Find—The consequences of underlying accident models in eight accident investigation manuals. *Saf. Sci.* **2009**, *47*, 1297–1311. [CrossRef]
41. Heinrich, H.W. *Industrial Accident Prevention: A Scientific Approach*; McGraw-Hill: New York, NY, USA, 1931.
42. Amalberti, R. The Paradoxes of almost totally safe transportation systems. *Saf. Sci.* **2001**, *37*, 109–126. [CrossRef]
43. Underwood, P.; Waterson, P. *Accident Analysis Models and Methods: Guidance for Safety Professionals*; Loughborough University: Loughborough, Leicestershire, UK, 2013.
44. Ibl, M.; Čapek, J. A Behavioural Analysis of Complexity in Socio-Technical Systems under Tension Modelled by Petri Nets. *Entropy* **2017**, *19*, 572. [CrossRef]
45. Carayon, P.; Hancock, P.; Leveson, N.; Noy, I.; Sznelwar, L.; van Hootegem, G. Advancing a sociotechnical systems approach to workplace safety–developing the conceptual framework. *Ergonomics* **2015**, *58*, 548–564. [CrossRef]
46. Smith, D.; Veitch, B.; Khan, F.; Taylor, R. Understanding industrial safety: Comparing Fault tree, Bayesian network, and FRAM approaches. *J. Loss Prev. Process Ind.* **2017**, *45*, 88–101. [CrossRef]
47. Patriarca, R.; Di Gravio, G.; Costantino, F. A Monte Carlo evolution of the Functional Resonance Analysis Method (FRAM) to assess performance variability in complex systems. *Saf. Sci.* **2017**, *91*, 49–60. [CrossRef]
48. Patriarca, R.; Falegnami, A.; Costantino, F.; Bilotta, F. Resilience engineering for socio-technical risk analysis: Application in neuro-surgery. *Reliab. Eng. Syst. Saf.* **2018**, *180*, 321–335. [CrossRef]
49. Mosleh, A. PRA: A Perspective on Strengths, Current Limitations, and Possible Improvements. *Nucl. Eng. Technol.* **2014**, *46*, 1–10. [CrossRef]
50. Paté-Cornell, E.; Dillon, R. Probabilistic risk analysis for the NASA space shuttle: A brief history and current work. *Reliab. Eng. Syst. Saf.* **2001**, *74*, 345–352. [CrossRef]
51. Wellock, T.R. A Figure of Merit: Quantifying the Probability of a Nuclear Reactor Accident. *Technol. Cult.* **2017**, *58*, 678–721. [CrossRef]
52. Kritzinger, D. *Aircraft System Safety: Assessments for Initial Airworthiness Certification*; Elsevier, Woodhead Publishing: Cambridge, UK, 2017. [CrossRef]
53. Zadeh, L.A. Is there a need for fuzzy logic? *Inf. Sci.* **2008**, *178*, 2751–2779. [CrossRef]
54. Leveson, N. The Use of Safety Cases in Certification and Regulation. *J. Syst. Saf.* **2011**, *47*, 1–9.
55. Alvarenga, M.A.B.; Frutuoso e Melo, P.F.; Fonseca, R.A. A critical review of methods and models for evaluating organizational factors in Human Reliability Analysis. *Prog. Nucl. Energy* **2014**, *75*, 25–41. [CrossRef]
56. Roy, N.; Eljack, F.; Jiménez-Gutiérrez, A.; Zhang, B.; Thiruvenkataswamy, P.; El-Halwagi, M.; Mannan, M.S. A review of safety indices for process design. *Curr. Opin. Chem. Eng.* **2016**, *14*, 42–48. [CrossRef]

57. Khan, F.I.; Husain, T.; Abbasi, S.A. Safety Weighted Hazard Index (SWeHI). *Process Saf. Environ. Prot.* **2001**, *79*, 65–80. [CrossRef]
58. Hollnagel, E. *Safety-I and Safety-II, The Past and Future of Safety Management*; Ashgate: Farnham, UK; Surrey, UK; Burlington, VT, USA, 2014; p. 200.
59. Leveson, N.G. Applying systems thinking to analyze and learn from events. *Saf. Sci.* **2011**, *49*, 55–64. [CrossRef]
60. Turner, B.A.; Pidgeon, N.F. *Man-Made Disasters*; Butterworth-Heinemann: Oxford, UK, 1978.
61. Perrow, C. *Normal Accidents: Living with High Risk Technologies*, 2nd ed.; Princeton University Press: Princeton, NJ, USA, 1984.
62. Le Coze, J.C. Reflecting on Jens Rasmussen's legacy. A strong program for a hard problem. *Saf. Sci.* **2015**, *71*, 123–141. [CrossRef]
63. Woods, D.D.; Cook, R.I. Nine Steps to Move Forward from Error. *Cogn. Technol. Work* **2002**, *4*, 137–144. [CrossRef]
64. Dekker, S.; Hollnagel, E.; Woods, D.D.; Cook, R. *Resilience Engineering: New Directions for Measuring and Maintaining Safety in Complex Systems*; Lund University School of Aviation: Lund, Sweden, 2008.
65. Borys, D.; Leggett, S. The fifth age of safety: The adaptive age? *J. Health Saf. Res. Pract.* **2009**, *1*, 19–27.
66. Xu, L.D.; He, W.; Li, S. Internet of Things in Industries: A Survey. *IEEE Trans. Ind. Inform.* **2014**, *10*, 2233–2243. [CrossRef]
67. Colgate, J.E.; Wannasuphoprasit, W.; Peshkin, M.A. Cobots: Robots for Collaboration with Human Operators. In Proceedings of the International Mechanical Engineering Congress and Exhibition, Atlanta, GA, USA, 17–22 November 1996; pp. 433–439.
68. Hollnagel, E. *FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-Technical Systems*; Ashgate Publishing, Limited: Farnham, UK, 2012.
69. Hollnagel, E.; Leonhardt, J.; Licu, T.; Shorrock, S. *From Safety-I to Safety-II: A White Paper*; Eurocontrol: Brussels, Belgium, 2013.
70. Hollnagel, E. Human Reliability Assessment in Context. *Nucl. Eng. Technol.* **2005**, *37*, 159–166.
71. Shorrock, S.T.; Kirwan, B. Development and application of a human error identification tool for air traffic control. *Appl. Ergon.* **2002**, *33*, 319–336. [CrossRef]
72. Boring, R.L.; Hendrickson, S.M.L.; Forester, J.A.; Tran, T.Q.; Lois, E. Issues in benchmarking human reliability analysis methods: A literature review. *Reliab. Eng. Syst. Saf.* **2010**, *95*, 591–605. [CrossRef]
73. Reason, J.; Hollnagel, E.; Paries, J. *Revisiting The Swiss Cheese Model of Accidents*; Eurocontrol Experimental Centre: Brétigny-sur-Orge, France, 2006.
74. Hollnagel, E.; Woods, D.D.; Leveson, N. *Resilience Engineering: Concepts and Precepts*; Ashgate: Farnham, UK, 2006.
75. Hale, A.; Borys, D. Working to rule, or working safely? Part 1: A state of the art review. *Saf. Sci.* **2013**, *55*, 207–221. [CrossRef]
76. Braithwaite, J.; Wears, R.L.; Hollnagel, E. *Resilient Health Care: Reconciling Work-as Imagined and Work-as-Done*; CRC Press: Boca Raton, FL, USA, 2017.
77. Hollnagel, E. The Nitty-Gritty of Human Factors. In *Human Factors and Ergonomics in Practice: Improving Performance and Well-Being in the Real World*; Shorrock, S., Williams, C., Eds.; CRC Press: Boca Raton, FL, USA; Taylor & Francis Group: Boca Raton, FL, USA, 2017.
78. Hollnagel, E. CREAM—Cognitive Reliability and Error Analysis Method. Available online: http://erikhollnagel.com/ideas/cream.html (accessed on 10 August 2018).
79. Bjerga, T.; Aven, T.; Zio, E. Uncertainty treatment in risk analysis of complex systems: The cases of STAMP and FRAM. *Reliab. Eng. Syst. Saf.* **2016**, *156*, 203–209. [CrossRef]
80. Underwood, P.; Waterson, P. Systemic accident analysis: examining the gap between research and practice. *Accid. Anal. Prev.* **2013**, *55*, 154–164. [CrossRef] [PubMed]
81. Patriarca, R.; Bergström, J.; Di Gravio, G. Defining the functional resonance analysis space: Combining Abstraction Hierarchy and FRAM. *Reliab. Eng. Syst. Saf.* **2017**, *165*, 34–46. [CrossRef]
82. Stanton, N.A. Representing distributed cognition in complex systems: How a submarine returns to periscope depth. *Ergonomics* **2014**, *57*, 403–418. [CrossRef] [PubMed]
83. de Vries, L.; Bligård, L.-O. Visualising safety: The potential for using sociotechnical systems models in prospective safety assessment and design. *Saf. Sci.* **2019**, *111*, 80–93. [CrossRef]

84. Leveson, N. A new accident model for engineering safer systems. *Saf. Sci.* **2004**, *42*, 237–270. [CrossRef]
85. Walker, G.H.; Gibson, H.; Stanton, N.A.; Baber, C.; Salmon, P.; Green, D. Event Analysis of Systemic Teamwork (EAST): A novel integration of ergonomics methods to analyse C4i activity. *Ergonomics* **2006**, *49*, 1345–1369. [CrossRef]
86. Salmon, L.; Stanton, N.A.; Walker, G.H.; Jenkins, D.P. *Distributed Situation Awareness: Theory, Measurement and Application to Teamwork*; Ashgate Publishing Limited: Farnham, UK, 2009.
87. Patriarca, R.; Bergström, J.; Di Gravio, G.; Costantino, F. Resilience engineering: Current status of the research and future challenges. *Saf. Sci.* **2018**, *102*, 79–100. [CrossRef]
88. Woltjer, R. Resilience Assessment Based on Models of Functional Resonance. In Proceedings of the 3rd Symposium on Resilience Engineering, Antibes-Juan-les-Pins, France, 28–30 October 2008.
89. Woods, D.D. Four concepts for resilience and the implications for the future of resilience engineering. *Reliab. Eng. Syst. Saf.* **2015**, *141*, 5–9. [CrossRef]
90. Moppett, I.K.; Shorrock, S.T. Working out wrong-side blocks. *Anaesthesia* **2018**, *73*, 407–420. [CrossRef]
91. Adriaensen, A.; Patriarca, R.; Smoker, A.; Bergstrom, J. A socio-technical analysis of functional properties in a joint cognitive system: A case study in an aircraft cockpit. *Ergonomics* **2019**. [CrossRef] [PubMed]
92. Patriarca, R.; Adriaensen, A.; Peters, M.; Putnam, J.; Constantino, F.; Di Gravio, G. Receipt and Dispatch of an Aircraft: A Functional Risk Analysis. In Proceedings of the 8th REA Symposium Embracing Resilience: Scaling Up and Speeding Up, Kalmar, Sweden, 24–27 June 2019.
93. Clay-Williams, R.; Hounsgaard, J.; Hollnagel, E. Where the rubber meets the road: Using FRAM to align work-as-imagined with work-as-done when implementing clinical guidelines. *Implement. Sci.* **2015**, *10*, 125. [CrossRef] [PubMed]
94. Green, B.N.; Johnson, C.D.; Adams, A. Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *J. Chiropr. Med.* **2006**, *5*, 101–117. [CrossRef]
95. Smith, K.M.; Valenta, A.L. Safety I to Safety II: A Paradigm Shift or More Work as Imagined? Comment on "False Dawns and New Horizons in Patient Safety Research and Practice". *Int. J. Health Policy Manag.* **2018**, *7*, 671–673. [CrossRef]
96. Nemeth, C.P.; Cook, R.I.; O'Connor, M.; Klock, P.A. Using Cognitive Artifacts to Understand Distributed Cognition. *IEEE Trans. Syst. Man Cybern.-Part A* **2004**, *34*, 726–735. [CrossRef]
97. Hollnagel, E.; Woods, D.D. *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*; Taylor & Francis: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2006.
98. Woods, D.D.; Roth, E.M. Cognitive Engineering: Human Problem Solving with Tools. *Hum. Factors* **1988**, *30*, 415–430. [CrossRef]
99. Bradshaw, J.M.; Hoffman, R.R.; Johnson, M.; Woods, D.D. The Seven Deadly Myths of "Autonomous Systems". *IEEE Intell. Syst.* **2013**, *28*, 54–61. [CrossRef]
100. de Winter, J.C.F.; Dodou, D. Why the Fitts list has persisted throughout the history of function allocation. *Cogn. Technol. Work* **2011**, *16*, 1–11. [CrossRef]
101. Dekker, S.W.A.; Woods, D.D. MABA-MABA or Abracadabra? Progress on Human-Automation Co-ordination. *Cogn. Technol. Work* **2002**, *4*, 240–244. [CrossRef]
102. Klein, G.; Woods, D.D.; Bradshaw, J.M.; Hoffman, R.R.; Feltovich, P.J. Ten Challenges for Making Automation a "Team Player" in Joint Human-Agent Activity. *IEEE Intell. Syst.* **2004**, *19*, 91–95. [CrossRef]
103. Hoffman, R.R.; Johnson, M.; Bradshaw, J.M.; Underbrink, A. Trust in Automation. *IEEE Intell. Syst.* **2013**, *28*, 84–88. [CrossRef]
104. Lyons, J.B.; Clark, M.A.; Wagner, A.R.; Schuelke, M.J. Certifiable Trust in Autonomous Systems: Making the Intractable Tangible. *AI Mag.* **2017**, *38*, 37–49. [CrossRef]
105. Nayyar, M.; Wagner, A.R. *When Should a Robot Apologize? Understanding How Timing Affects Human-Robot Trust Repair*; Springer: Cham, Germany, 2018; pp. 265–274.
106. D'Addona, D.M.; Bracco, F.; Bettoni, A.; Nishino, N.; Carpanzano, E.; Bruzzone, A.A. Adaptive automation and human factors in manufacturing: An experimental assessment for a cognitive approach. *CIRP Ann.* **2018**, *67*, 455–458. [CrossRef]
107. BEA. *Final Accident Report on Flight AF 447*; Accident Report; Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile: Paris, France, 2012.
108. Nederlandse Onderzoeksraad voor Veiligheid. *Neergestort Tijdens Nadering, Boeing 737–800, Nabij Amsterdam Schiphol Airport, 25 Februari 2009*; Nederlandse Onderzoeksraad voor Veiligheid: The Hague, The Netherlands, 2010.

109. Sarter, N. Investigating mode errors on automated flight decks: Illustrating the problem-driven, cumulative, and interdisciplinary nature of human factors research. *Hum. Factors* **2008**, *50*, 506–510. [CrossRef] [PubMed]
110. Sarter, N.B.; Woods, D.D. How in the World Did We Ever Get into That Mode? Mode Error and Awareness in Supervisory Control. *Hum. Factors* **1995**, *37*, 5–19. [CrossRef]
111. Dekker, S. *Patient Safety: A Human Factors Approach*; CRC Press: Boca Raton, FL, USA; London, UK; New York, NY, USA; Francis and Taylor Group: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2011.
112. Woods, D.D. The Risks of Autonomy. *J. Cogn. Eng. Decis. Mak.* **2016**, *10*, 131–133. [CrossRef]
113. Wiener, E.L. *Human Factors of Advanced Technology ("Glass Cockpit") Transport Aircraft*; Ames Reseacrh Center, NASA: Moffett Field, CA, USA, 1989.
114. Cook, R.I.; Woods, D.D.; Howie, M.B.; McColligan, E. Cognitive High Workload, Consequences Of High Consequence 'Clumsy' Automation Human Performance. Paper presented at the Fourth Annual Workshop on Space Operations Applications and Research (SOAR 90), Albuquerque, NM, USA, 26–28 June 1990; pp. 543–546.
115. Lee, J.D.; Seppelt, B.D. Human Factors and Ergonomics in Automation Design. In *Handbook of Human Factors and Ergonomics*; Salvendy, G., Ed.; Wiley: Hoboken, NJ, USA, 2012. [CrossRef]
116. Woods, D.D.; McColligan, E.; Howie, M.B.; Cook, R.I. *Cognitive Consequences of Clumsy Automation on High Workload, High Consequence Human Performance*; NASA Publication: Washington, DC, USA, 1991.
117. Fitts, P.M. *Human Engineering for an Effective Air-Navigation and Traffic-Control System*; National Research Council: Washington, DC, USA, 1951.
118. Bainbridge, L. Ironies of automation. *Automatica* **1983**, *19*, 775–779. [CrossRef]
119. Woods, D.D.; Sarter, N.B. Capturing the dynamics of attention control from individual to distributed systems: The shape of models to come. *Theor. Issues Ergon. Sci.* **2010**, *11*, 7–28. [CrossRef]
120. Department of Defense. *Autonomy Community of Interest (COI), Test and Evaluation, Verification and Validation (TEVV) Working Group, Technology Investment Strategy 2015–2018*; Office of the Assistant Secretary of Defense for Research & Engineering: Washington, DC, USA, 2015.
121. Dekker, S. *Report of the Flight Crew Human Factors Investigation Conducted for the Dutch Safety Board Into the Accident of TK1951, Boeing 737–800 Near Amsterdam Schiphol Airport, February 25 2009*; Lunds Universitet, School of Aviation: Lund, Sweden, 2009.
122. Banks, V.A.; Stanton, N.A. Analysis of driver roles: Modelling the changing role of the driver in automated driving systems using EAST. *Theor. Issues Ergon. Sci.* **2019**, *20*, 284–300. [CrossRef]
123. Dekker, S. *The Field Guide to Understanding Human Error*; Ashgate: Farnham, UK, 2006.
124. Abdulkhaleq, A.; Wagner, S.; Leveson, N. A Comprehensive Safety Engineering Approach for Software-Intensive Systems Based on STPA. *Procedia Eng.* **2015**, *128*, 2–11. [CrossRef]
125. Friedberg, I.; McLaughlin, K.; Smith, P.; Laverty, D.; Sezer, S. STPA-SafeSec: Safety and security analysis for cyber-physical systems. *J. Inf. Secur. Appl.* **2017**, *34*, 183–196. [CrossRef]
126. Patriarca, R.; Di Gravio, G.; Costantino, F. myFRAM: An open tool support for the functional resonance analysis method. In Proceedings of the 2017 2nd International Conference on System Reliability and Safety (ICSRS), Milan, Italy, 20–22 December 2017; pp. 439–443.
127. Albery, S.; Borys, D.; Tepe, S. Advantages for risk assessment: Evaluating learnings from question sets inspired by the FRAM and the risk matrix in a manufacturing environment. *Saf. Sci.* **2016**, *89*, 180–189. [CrossRef]
128. Gattola, V.; Patriarca, R.; Tomasi, G.; Tronci, M. Functional resonance in industrial operations: A case study in a manufacturing plant. *IFAC-PapersOnLine* **2018**, *51*, 927–932. [CrossRef]
129. Melanson, A.; Nadeau, S. *Managing OHS in Complex and Unpredictable Manufacturing Systems: Can FRAM Bring Agility?* Springer: Cham, Germany, 2016; Volume 490, pp. 341–348.
130. Zheng, Z.; Tian, J.; Zhao, T. Refining operation guidelines with model-checking-aided FRAM to improve manufacturing processes: A case study for aeroengine blade forging. *Cogn. Technol. Work* **2016**, *18*, 777–791. [CrossRef]
131. Salmon, P.M.; Read, G.J.M.; Walker, G.H.; Goode, N.; Grant, E.; Dallat, C.; Carden, T.; Naweed, A.; Stanton, N.A. STAMP goes EAST: Integrating systems ergonomics methods for the analysis of railway level crossing safety management. *Saf. Sci.* **2018**, *110*, 31–46. [CrossRef]
132. Patriarca, R.; Falegnami, A.; Bilotta, F. Embracing simplexity: The role of artificial intelligence in peri-procedural medical safety. *Expert Rev. Med. Devices* **2019**, *16*, 77–79. [CrossRef]

133. Abdulkhaleq, A.; Wagner, S. *XSTAMPP: An eXtensible STAMP Platform as Tool Support for Safety Engineering*; University of Stuttgart: Stuttgart, Germany, 2015.
134. Patriarca, R.; Constantino, F.; Di Gravio, G. myFRAM: An IT Open Tool to Support the Application of the FRAM. Available online: https://functionalresonance.com/onewebmedia/23%20FRAMily%202018_myFRAM_Patriarca%20Riccardo%20et%20al.pdf (accessed on 10 August 2019).
135. The Functional Resonance Analysis Method, FRAM Model Visualiser (FMV). Available online: https://functionalresonance.com/FMV/index.html (accessed on 15 August 2019).
136. Patriarca, R.; Del Pinto, G.; Di Gravio, G.; Costantino, F. FRAM for Systemic Accident Analysis: A Matrix Representation of Functional Resonance. *Int. J. Reliab. Qual. Saf. Eng.* **2017**. [CrossRef]
137. Or, L.B.; Arogeti, S.; Hartmann, D. Challenges in Future Mathematical Modelling of Hierarchical Functional Safety Control Structures within STAMP Safety Model. In Proceedings of the MATEC Web of Conferences, Amsterdam, The Netherlands, 2 November 2018. [CrossRef]
138. Rosa, L.V.; Haddad, A.N.; de Carvalho, P.V.R. Assessing risk in sustainable construction using the Functional Resonance Analysis Method (FRAM). *Cogn. Technol. Work* **2015**, *17*, 559–573. [CrossRef]
139. Vincent, C.; Amalberti, R. *Safer Healthcare, Strategies for the Real World*; Springer: Berlin/Heidelberg, Germany, 2016.