Reliability Engineering and System Safety 210 (2021) 107547

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Debates and politics in safety science

Torgeir Kolstø Haavik

NTNU Social Research, Dragvoll Allé 38 B, 7491 Trondheim, Norway

ARTICLE INFO

Keywords:
Safety science
Complexity
Uncertainty
Politics

ABSTRACT

This article explores what complementary perspectives Science and Technology Studies and in particular Actor Network Theory may bring to safety science beyond what comes out of traditional comparisons between highly profiled theories/perspectives of Normal Accident Theory (NAT), High Reliability Organisations (HRO), Resilience Engineering (RE) and Safety II.

In the article, core ideas of NAT, HRO and RE/Safety II are reviewed, and debates over NAT/HRO, HRO/RE and Safety I/Safety II are discussed. Thereafter, controversies over complexity, non-events and uncertainty respectively are identified and elaborated, drawing on a richer repertoire from the social sciences, in particular Actor Network Theory. The article concludes by inviting to more serious engagement in scientific controversies and politics of safety, and operationalises this into three propositions: Take complexity seriously; broaden the perspectives and methodologies for understanding sociotechnical work; and make safety science research politically oriented (again).

1. Introduction

Safety research as a systematic, scientific subject is fairly young, with the pioneer works of the social science and organisational approaches to safety dating back to the seventies (e.g. [80]). Since then, safety research has developed through different phases, each marked with different focus areas. Hence one often refers to the three ages of safety, each characterised by different foci and different types of attributed causes for accidents and different research scopes[17]. The first, technological age of safety was followed by the age of human factors, after which safety research entered the age of organisational attention including such themes as safety culture and safety management systems. It is particularly after entering the age of organisations that the social science approach to safety really expanded both in terms of volume and perspectives.[1] This development was fuelled by a number of organisational accidents during the 1980s; Bhopal[7], Chernobyl[66], Three Mile Island[60], Piper Alpha(Pat [59] and Challenger[83] are all examples of accidents that spurred substantial research activities with the aim not only to establish the causes of the accidents, but also to develop theories of risk and safety that reflected the developments that had taken place in terms of increased complexity of industrial sociotechnical systems.

It was in this period that the social sciences seriously entered the arena of safety research. Surely, the social sciences was already represented, Barry Turner [80–82] and Charles Perrow [60] – both sociologists – had at this stage already contributed with strong empirically informed theoretical frameworks that still enjoy a central position in the field of safety research, and in light of that they should be considered pioneers in the social studies of risk. Until the 1980s, the field of safety research had largely been dominated by contributions from engineering and psychology, and foundational contributions from Reason [67,68] and Rasmussen [64,65] have kept on gaining momentum and influencing the thoughts of safety scholars up till today. However, from the 1980s onward, safety science increasingly attracted scholars with background from sociology, anthropology, organization and management science. Examples are the sociologist Diane Vaughan with her ethnographic studies in NASA following the Challenger accident[83], sociologist Andrew Hopkins' study of the gas plant explosion at Longford [28], the organizational theorist Karl Weick with his works on organisational sensemaking[89], and the Berkeley research group and their work on High Reliability Organisations (HRO) [40,41,90–92].

Among the most recent contributions to the field of safety research that has achieved status as a school of its own is Resilience Engineering [23–25, 27,56], developed and nurtured by Erik Hollnagel, David Woods, Sidney Dekker, Jeffrey Braithwaite, Jean Pariès and others,[2]

E-mail address: torgeir.haavik@samforsk.no.

[1] The present description of the development of safety science research is not intended to be exhaustive. See e.g. Le Coze [54] or Dekker [13] for a thorough review.

[2] Surely, the mentioning of names immediately calls attention to others that are omitted. For example, the roots of Resilience Engineering can be traced back to both Jens Rasmussen and James Reason (see [55] for a review of the history of Resilience Engineering).

https://doi.org/10.1016/j.ress.2021.107547
Received 27 January 2020; Received in revised form 2 February 2021; Accepted 10 February 2021
Available online 18 February 2021
0951-8320/

with the derivative concepts of Safety II (and Safety I), contrasting conceptually and qualitatively different approaches to safety [14,20,22,26,79].

Along with the increased research activity and increased public interest in safety, debates between different perspectives became important for the nurturing and scaffolding of research communities that subscribed to the different perspectives. In this paper I shall look into three such debates; that between Normal Accident Theory and High Reliability Organisations, that between High Reliability Organisations and Resilience Engineering, and the debate over Safety I/Safety II. The article summarises these debates as they appear as different orientations and traditions to safety.

The core arguments of these debates is well known matter, and they have been remarkably stable over the years, indicating that there are still unresolved issues there. On the other hand, much criticism has been raised about these debates, about the lack of precise definition of core concepts and whether there really are significant and sufficiently defined differences between the foundational assumptions in the respective traditions: that allowed or conceivable variations in behavior of systems is insufficiently accounted for when referring to accidents as normal or abnormal; that the role of public visibility of high reliability organisations is neglected, thus also how normal accidents return as public attention decreases; that success is insufficiently defined, and when boundaries for success are defined Safety II reverts to Safety I; that RE has so many similarities to HRO that it is hard to see that it represents something new; that when people continue to make reference to NAT, it is not because that work supports their arguments but simply to establish that they are aware of the relevant literature; and that research/science and business are so intertwined that there is an additional rationale for advertising scientific differences (see e.g. [2–4, 29,58]).

I do not take a stance to this criticism here, but instead I ask: What if there are more and other things to these debates than has so far been explicitly exposed? What if there are some more foundational issues at stake, that can be excavated from these discourses? The article pursues this challenge by discussing how the debates hide fundamentally different perspectives on *complexity, events* and *uncertainty*. In this discussion, the article draws on literature from outside the traditional safety library, in particular writings on workplace studies (e.g. [76,78]) and Actor Network Theory (e.g. [18,46]). The consequence of the analysis is operationalised into three propositions for future safety science research.

In the article I shall develop a richer understanding of these central debates that are played out in the research literature. Central journals including Safety Science and Journal of Contingencies and Crisis Management, in which there are also special issues to be found ([39,50,88]), explicitly facilitate the debates. Other, less systematic facilitation for the debates can also be found in the literature, and I shall adopt the pragmatic approach of accounting for them where they appear most visible.

The paper is structured such that first, historic and contemporary debates within the safety science community are presented. I label these debates overexposed, certainly not with reference to the quality of the debates or the arguments that are held, but due to their persistence despite a lack of progress and a tendency to stabilize and stall at a very early stage. Thereafter, I shall, following a mode of reasoning inspired by Actor Network Theory, lend words to some tacit themes that claim higher criticality than the core issues of the 'standard' debates. I label these themes underexposed, certainly not with reference to any sort of scientific superiority, but because they represent turns that the safety science discourse may take if it is opened up again and reach out from the sometimes self-referential scientific field. But before I proceed with this, I shall briefly account for the theoretical perspectives that engage into these debates, and the justification for bringing perspectives from Science and Technology Studies into safety science research.

1.1. Highly profiled theories of organisational safety and risk

Three approaches to safety have acquired a central place in safety science research, due to the combination of two achievements – that of formulating a foundational theory of safety (or why some organisations fail and others succeed), and that of constructing or taking actively part in debates where the different theories or perspectives are put up against each other; while many pioneers have a share in what was to become a recognizable field of safety science from the 1980ies and onward, the perspectives of Normal Accident Theory, High Reliability Organisations and Resilience Engineering have a special status with respect to prevalence and attention. Their core messages as widely known by now:

According to Normal Accident Theory systems with tight coupling – with tightly prescribed steps and invariant sequences that cannot be changed, implying that interruptions propagate rapidly through the system, and that there is little slack and little room for improvisation – require a control mode that is centralised. Systems with complex interactions – where accidents tend to stem from "mysterious interactions of failures" ([61], 10) – require a decentralised control modus. These requirements are sometimes sources for organisational contradictions; since sociotechnical systems cannot be centralised and decentralised at the same time, systems that are both tightly coupled and interactively complex cannot in the long run be managed in a safe manner.

High Reliability Organisations are characterised by two distinct features: organisational redundancy – both structural and cultural – and the ability to reconfigure spontaneously and more or less seamlessly rearrange from a modus of centralised control to a management mode of decentralised control when experiencing a crisis and going from a normal work mode to a mode of crisis management [40,41]. Another description of HROs are their five characteristics of collective mindfulness – *preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience*, and *underspecification of structures* [93].

Resilience Engineering is inspired by the ecological references to resilience, and if one is to speak of the essence of resilience, itis "the intrinsic ability of an organization (system) to maintain or regain a dynamically stable state, which allows it to continue operations after a major mishap and/or in the presence of a continuous stress" ([27], 16). As RE has its roots in cybernetics thinking, central themes are *adaptation, variability* and *functional resonance*. The recommendation of RE to study and learn from that which goes right has given rise to an alternative definition of safety – *Safety II* – the "system's ability to succeed under varying conditions" ([26], 4). The counterpart to Safety II is labelled *Safety I*, defining safety as the absence of undesirable events and accidents, and freedom of unacceptable risk.

1.2. Science and Technology studies perspectives in safety science research

As with Safety science, one might say that Science and Technology Studies (STS) is a young academic field. Growing out of the early orientations towards the sociology of knowledge (e.g. [10,38]), it manifested itself strongly with the early contributions from e.g. Latour and Woolgar [47], Pinch and Bijker [63] and Callon [12], some of who were also central in the development of Actor Network Theory (ANT) [43,46] – a branch of STS. Since then, STS perspectives has become naturalised in many of the more established disciplines and fields, for example sociology, geography, and computer science.

STS perspectives and methods have not been completely absent in safety science, but it would be an exaggeration to say that they have been numerous and very visible. One prominent example, however, is the works of Diane Vaughan (e.g. [83–85]), where the relevance of STS for safety research is convincingly demonstrated. Other authors who have successfully adopted STS perspectives in their writings include – but are surely not limited to – Le Coze [53], Almklov and Antonsen [5,57]. Personally, I have found significant resonance in the methodologies and the mode of reasoning in Actor Network Theory (ANT). The liquid characteristic of ANT lending it a status located somewhere in the intersection between ontology, theory and methodology of phenomena where the social and the material is tightly coupled, makes it useful in the exploration of sociotechnical systems and sociotechnical work in a context of risk and safety [30,32–35].

The way established issues and themes are treated in ANT works is often different – and sometimes surprising – from the common treatments, and not seldom this is done with a critical and renewed view on what analytical concepts represent and what empirical findings are really cases of – as for example in studies of complexity [77], modes and methods of scientific work [42,44] and the ontology of risk and uncertainty [18,45,46,48]. Besides alternative and often constructivist approaches to established subjects, the ANT catalog offers a substantial repertoire and track record on empirical research methods particularly on social aspects of technology, and material aspects of the social. It has always been an ambition for safety research to traverse this bridge, safety science indeed having been an interdisciplinary field from the very beginning, but the field as such is still struggling to be able to stand safely with feet both in the social tradition and in the engineering tradition simultaneously. ANT was developed with exactly such an ambition, and has much to offer safety science in this respect.

ANT research is always looking for controversies, as they represent cracks where light gets in and makes visible what phenomena/systems are made up of.[3] Such controversies need not be major – they may not even have to be explicated, and may thus have to be actively searched for. In the following I shall do exactly that; heat up some controversies that may lie hidden in the shadows of some less heated debates in safety science. The aim is to inspire to creative search for doors to new rooms in safety science research, where we may find new questions and answers relevant for the old world of industrial and organisational safety, but also for the new world of societal resilience that is being gestalted as we speak.

2. Overexposed debates in safety science

NAT, HRO and RE can be considered as canonical perspectives in the organisational branch of safety science. Representing different 'schools', they are nurtured and developed in different research communities with little overlap. It could be said that they are living peacefully side by side, but although as a general judgment that would be correct, it conceals some important nuances: between NAT and HRO there has been occasions of explicit controversies with considerable temperature; HRO and RE has been 'accused' of being without any significant difference, and still there are debates around the fundamental approach to safety; the Safety II perspective is at odds with the Safety I perspective, but does not really have a well-defined discussion partner since Safety I is a construct originating from Safety II advocate Erik Hollnagel and rarely used as a label by anyone outside the 'Safety II community'.

It should be noted that the three debates that are presented are by no means symmetrical with respect to their position within, not to say beyond, the safety science community. In its time, the NAT/HRO debate extended beyond the community of safety professionals and engaged prominent US sociologists and political scientists, and involved tangible scientific debates, some of which took place in the Journal of Contingencies and Crisis Management (e.g. La [39]). The HRO/RE debate, on the other hand, can also be called a non-debate, as the relation between the two is apparently in most respects characterised by similarities than by differences (see e.g. [29]). Still, and despite remarkably little interaction and communication between the fields, also this admittedly more safety science-internal debate has taken place in dedicated journal issues (e.g. [88]). The third debate presented here, the Safety I/Safety II debate, is even more internal within the safety science community and actually with only one party – those advocating for Safety II. This is not to say that it is a less important debate, and this comparison between debates is not an attempt to range them in any way – indeed, any debate takes time to mature and grow in intensity and prevalence, and the Safety I/Safety II debate is very young, so is the HRO/RE debate, while the NAT/HRO debate has a long history and took some time to develop – but it should be kept in mind that the debates selected here are selected idiosyncratically (though on the basis of visibility and volume of publications in the safety science catalog), without reference to established discourses on central debates in safety science research.

2.1. The NAT/HRO debate

While NAT and HRO are oriented towards the same type of sociotechnical systems, they come to different conclusion with respect to the prospects of system safety in the long run. The basic message in NAT is that accidents are inevitable in complex, tightly coupled sociotechnical systems because the combination of unexpected interaction and rapid propagation and escalation confronts the management with an insoluble control paradox [60]. According to HRO, this conclusion does not hold water; empirical studies of highly reliable organisations have shown that HROs adopt organisational strategies where they centralize the design of decision premises in order to allow decentralized decision making. In this manner the required line of operating with both centralised and decentralised control in parallel is no longer a paradox [70,72,93].

In addition, the evaluation of redundancy differs between NAT and HRO. While HRO promotes redundancy because "if one component fails, another backs it up; if one operator fails to carry out his task, another one takes over his position; if danger lurks, multiple channels are used to transmit warnings" [70], NAT has a more ambivalent relation to redundancy since redundant components may depend on a common denominator. This was the case with the Challenger space shuttle's Solid Rocket Booster's sealing; the famous O-ring that failed during the Challenger launch was actually backed up with another O-ring. However, both O-rings were dependent on weather conditions – and both failed during the launch[83].

2.2. The HRO/RE debate

With the introduction of Resilience Engineering came also objections that RE did not represent something new. HRO had been occupied with resilience for a long time already, and just like RE researchers, HRO researchers had spent much time on understanding why things go right; the very origin of the HRO research was based on studies of well-performing organisations in accident-prone environments. A critique against RE is thus that "it offers itself as something new, when in fact it is hard to see in what way its 'precepts and concepts' depart from those of HRO theory" ([29], 9). This is a critique that receives little explicit opposition in terms of factual arguments, and in light of that the RE research literature the last decade with little or no reference to HRO literature is surprisingly voluminous.[4] The HRO/RE debate is therefore first and foremost a silent debate with few or no cross-references [29], for which reasons may very well be highly pragmatic [36].

There are more explicit disagreements between the two perspectives, however, both with respect to whether adaptive capacities should be considered as an individual (RE) or a cultural (HRO) character ([62], 4), and with respect to treating resilience as a generic capacity (RE) contrasted with the importance of "[parsing] out and empirically ground

[3] Alternatives to studies of controversies are studies of breakdowns, which offer the same kind of inspection possibilities. This is an obvious rationale of accident investigations, which are so fundamental in the construction of safe systems.

[4] For an overview, see e.g. [9,69]. The number of books alone on Resilience Engineering also say something about the momentum of the theory [11,21–24, 27,56,87]

[ing] the concept of resilience into different types"([62], 2) of *precursor* resilience, *restoration* resilience and *recovery* resilience. However, these disagreements are not subject to either much attention or particularly loud debating.

2.3. The Safety I/Safety II debate

With the introduction of a new safety perspective – Safety II [21] – came also a new label for much of the safety research that Safety II criticised – 'Safety I'. While there are few or no references to Safety I among those who are given that label, there are neither many objections to the type of categorization that was introduced with the terms. The Safety I/Safety II debate is thus a dawning, but still a low-intensity debate, addressing two issues: whether we learn more from failures or from successes, and whether or not safe operations are actually visible and possible to study.

The practice of learning from failures has a long tradition in safety science,[5] and has been foundational for the incremental development of aviation safety, for example.

The argument for the better learning potential within the Safety II paradigm is the asymmetric occurrence of intended and unintended outcomes. This argument lends validity from the theory that unintended and intended outcomes have the same genealogy: namely the natural occurring variability and (sometimes) resonance; "Resilience engineering acknowledges that acceptable outcomes and unacceptable outcomes have a common basis, namely everyday performance adjustments" ([26], 16).

The argument of visibility and invisibility is already mentioned in connection with the perspective on safety as a dynamic non-event. Hollnagel gives credit to Weick for this definition of safety, but still finds it problematic. Here, the argument for Safety II is an argument of measurement and *dynamic events*;

"Even when people agree that safety is a dynamic-non-event, the practice of safety management is to count the events, i.e., the number of accidents, incidents and so forth. By doing that we know how many events there have been, but not how many non-events. We may, however, easily turn the tables, by defining safety as a dynamic *event.*" The event is now an activity that succeeds or goes well (that we come home safely, that the plane lands on time, etc.), and we are obviously safe when that happens. The non-event consequently becomes the situation when this does *not* happen, i.e., when things go wrong. We can count the non-events, i.e., the non-successes or failures, just as we have usually done. But we can now also count the events, the number of things that go right, at least if we make the effort." ([21], 9)

The magnitude of the effort we need to make, however, should not be underestimated. However, it is not only a question of effort, but also of *method*. To see more examples in the future of thorough accounts of work in terms of *events*, it is not enough to allocate resources for counting; we would also need to develop the *methods* of describing work. More about that below.

From this review of debates within a safety science frame of reference, we shall now shift our position and reflect upon these safety discourses from a different angle, drawing on a richer repertoire from the social sciences.

3. Underexposed themes in safety science

The debates reviewed above have different historic trajectories, different ages and appeal to different participants, but they all share the fate of struggling with progression. What can Science and Technology Studies offer safety science in its search for matters of concern that can bring theory building further? Inspired by STS' interest in epistemology and controversies, we may discover matters of concern that we do see as clearly or do not have a sufficient vocabulary for when studying them within a safety theory framework. The following sections explore this opportunity. The themes discussed below should be considered as a few examples among many, and they are probably not the most essential either. They should be considered as examples, that at best can inspire further and more systematic inventory of themes in future works.

3.1. The natures of complexity

A general and explicit conception of complexity in safety science, and across NAT, HRO and RE, is that complexity makes systems more accident-prone. Complexity is thus generally seen as a negative entity, and something that should be minimized. Seen from an epistemological perspective, there are two noteworthy issues that have not gained much attention in safety science. First, if we look into the writings on complexity in NAT and RE respectively, we find different ways to relate to complexity that are seldom made explicit in the safety science discourses. Second, views on complexity that include the *social* are seldom discussed, although we deal mostly with sociotechnical systems, and I shall here make a comment on alternative definitions of complexity that may shed new light on existing safety science debates. Importantly, although Perrow's and Hollnagel's approaches to complexity are those that are discussed here, I certainly do not imply that these approaches and authors are the only sources to writings on complexity in safety science; there are many other valuable sources, including Vaughan [84] and [49,51,52].

In Perrow's writings, complexity refers to "interactions in an unexpected sequence" ([61], 78). It is underscored that "systems are not (…) complex, strictly speaking, only their interactions are" ([61], 78).[6] Still, complexity is seen as static, to the extent that the complexity of different systems may be plotted on a scale. Hollnagel [19] refers to the same phenomenon as Perrow when he writes about complexity, but he is more explicit on the polysemantics of complexity. Several formal and informal definitions are suggested; from complexity referring to systems involving many parts, to Wiener's [94] definition that a system's complexity reflects the degree of difficulty in predicting the properties of the system if the properties of the system's parts are given.

So far, complexity is largely treated as an ontological entity in the positivist tradition. It is first when Hollnagel introduces *intractability* we understand that Perrow and Hollnagel are actually talking about two different kinds of complexity. In Hollnagel's words,

"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." ([19], 204)

If we adopt a social science perspective on these research strands, we may say that Perrow (NAT) talks about *positivist ontological complexity*, and Hollnagel (RE) talks about *epistemological complexity*, or a constructivist ontological complexity. The latter type of complexity lies much closer to the type of complexity that[77] portray in their study of social life among baboons; however, here the degree of complexity is inversely proportional with the wealth of material and immaterial tools available for organizing work and social relations:

"When [baboons] construct and repair their social order, they do so only with limited resources, their bodies, their social skills and whatever social strategies they can construct." ([77], 790)

The different notions of complexity are seldom explicated in the safety science literature, and we may ask: how shall we understand that complexity is such a central term in safety science, and that these different meanings exist but are not being seriously tackled in safety science? If we take seriously Strum and Latour's definition of sociotechnical complexity, we might come closer to the dirty reality of

[5] See e.g. Drupsteen and Guldenmund [14], Stanton et al. [75].

[6] For later reference: without stating it explicitly Perrow here hints towards the relational nature of complexity.

sociotechnical systems, and the Janus face of complexity: sociotechnical systems are never intrinsically resilient, and their complexity is never static. It can change all of a sudden. When something breaks, Pandora's box opens and complexity suddenly rises.

This leads to the first proposition of this paper: Stimulate and enrich safety science discourses on complexity of sociotechnical systems with new perspectives from the social sciences that have experience with addressing social (aspects of) complexity.

3.2. The richness of events

As we have seen already, discussions on visibility and auditability enter into the debates of safety science. Weick's [90] referencing to safety as a non-event associates safety with invisibility, while Hollnagel's [21] referencing to safety as *events* introduces the idea of counting positive outcomes for understanding the intangibility of the Safety I/Safety II debate; while Weick refers to the difficulty of accounting for non-events that do *not* happen (accidents), Hollnagel argues for counting the things that *do* happen – the desired outcomes. In lack of a common reference, the Safety I/Safety II debate is not configured to be resolved, it may seem, but carries a potential for connecting to discourses outside the field of safety science, such as those on articulation work and invisible work in the workplace studies, computer-supported cooperative work and Science and Technology Studies (see e.g. [15] for an appreciation of how these fields overlap and merge into something we might call sociology of work).

Beyond the question in safety science of *what* to look for is the question of *how* to do it. What is visible and what is invisible is not a question of essence, but of method. Although one might think that the question of *how* could actually solve the Gordian knot of Safety I/Safety II, it is not much addressed in safety science research. If we look to workplace studies, work that is carried out, particularly in the face of contingencies, but is not seriously acknowledged because it goes 'below the radar', is sometimes analysed as articulation work. There is a rich literature on articulation work that provides theories, methods and tools for accounting for articulation work and 'invisible work' (e.g. [76,78], 1995, Strauss 1985, Gerson and Star 1986, Schmidt and Bannon 1992, Bannon and Bødker 1997, Haavik 2010). This literature offers rich empirical sociomaterial descriptions of 'dynamic events', although with a vocabulary that may not always translate immediately to the language and jargon of safety science.

This leads to the second proposition of this paper: Look to the wider sociology of work for inspiration and experience with empirical studies of work in high-risk sociotechnical systems. Whether the results from such endeavours feed into discussions on definitions of safety as dynamic non-events or dynamic events, produce new constructs such as sensework [33,34] or renew safety science orientation in other ways is not decisive, as long as it deepens our discussions and brings safety science and research forward. There is no guarantee that this will happen, but it can be worthwhile to try.

3.3. The politics of risk and uncertainty

While risk has traditionally been defined as probability*consequence, an increasingly used definition of risk in the later years is that "[*r*]isk refers to uncertainty about and severity of the consequences (or outcomes) of an activity with respect to something that humans value"([6], 2). In connection with risk management, defined situations of hazard and accident (DSHA) [74,86,95] represent the activities – or the risk objects – for which risk and uncertainty are addressed. While DSHA is central in the science and practice of safety, it is at odds with the strong advice from sociologist and STS scholar Stephen Hilgartner. Hilgartner advises us to study risks objects not as ready-made, but as objects in the making: "Perceptions of risk are not things that get tacked onto technology at the end of the day. Definitions of risk get built into technology and shape its evolution" ([18], 39). The tension between ready-made and in-the-making approaches in science in general has been thoroughly addressed by Latour [43], while [18] has addressed this relation in particular in connection with risk objects. Both Latour and Hilgartner seek to avoid black-boxed notions of risk. In Latour's [45] reading of Beck [8], "[risk] does not mean that we run more dangers than before, but that we are now *entangled*. (…) A perfect translation of 'risk' is the word *network* in the ANT[7] sense, referring to whatever deviates from the straight path of reason and of control" ([45], 36). And what Latour [46] does in *Reassembling the social* is to systematize a methodology for how such sociotechnical networks can be studied in the social sciences. Hilgartner builds on Latour and shows how risk objects in the motor vehicle domain are constructed and distributed over sociotechnical networks:

"Thus, the risk of motor vehicle accidents can be attributed to unsafe drivers, unsafe roads, or unsafe cars [16,37]. The risks posed by unsafe drivers, in turn, can be attributed to inexperience, irresponsibility, fatigue, or alcohol consumption. Inexperience, in turn, can be attributed to inadequate driver education programs, which can be attributed to shortages of funds. The networks of objects deemed to pose risks can abruptly terminate, can wind their way back to common causes, or can diverge. The set of possibilities, in principle, is infinite." ([18], 42)

While this framing of risk is not very different from "underlying causes" in safety science, Hilgartner goes further to shed light on the *politics of risk*; the process of negotiating what are to be considered risk objects. Hilgartner refers to the work of Gusfield:

"Gusfield [16] has shown how the 'killer drunk' was singled out as *the* cause of traffic fatalities that involve alcohol. As he argued, this way of framing the problem allows car makers and the alcoholic beverage industry to 'disown' the problem, shifting causal, moral and political responsibility for traffic deaths onto other actors. Directing attention at the individual drunk driver as the risk object deflects attention from alternative risk objects–such as unsafe cars, or land use and transportation policies that permit bars to be sited in locations accessible only by automobile."([18], 43)

Considering, for example, the developments of digitalisation in the petroleum industry the last decade and the new division of labor between onshore and offshore – an operations regime known as Integrated Operations[1,31,73] – one would anticipate a whole new process of negotiating risk objects. Still, a study of the impact of Integrated Operations on risk objects – defined situations of hazard and accident – conclude that "there is currently no need for changing the DSHAs currently used in the industry, as the overall types of accidents that may arise in an IO environment are foreseen to be similar to the types of accidents that may arise in a traditional operational environment"([74], 2). If we consider this approach to DSHA and compare it to Latour's and Hilgartner's references to uncertainty, networks and the social construction of risk objects,[8] we are right in the middle of an argument over risk as a ready-made object, and risk as in-the-making, or between *essentialistic* risk and risk as a *socio-technical-political construct* respectively. We are of course well familiar with this debate on uncertainty (and risk), where external and internal (objective and subjective) notions of uncertainty also frequently blend in, but still we rarely see the political aspects of risk addressed in safety research. Actually, this political aspect is one of the most central conclusions in Perrow's [60] work on normal accidents, and since, but still the one taken least serious. Maybe his scepticism towards certain sociotechnical configurations came too early, before we had seen the mechanisms and effects of climate change, just to mention one example, for real.

The 'pessimist' advice has strong political implications, and need stronger argumentation than what can be provided from within safety science frameworks alone. However, Latour's and Hilgartner's constructivist approaches to risk and uncertainty offer a different

[7] Actor Network Theory. See e.g. Latour [46].

[8] Which, when we look inside them, are sociotechnical networks.

framing that mayrevitalize the old NAT/HRO debate. For example, the contemporary discourse on societal safety is almost exclusively oriented towards making critical infrastructures and institutions reliable and robust, such as energy infrastructures and (increasingly) energy-consuming individual and societal life practices. But what if it is exactly these infrastructures and institutions that produces the risks we are currently trying to fight? Then we should not only talk about vulnerable infrastructures, but also of *infrastructures of risk*.

This leads to the third proposition of this paper: make safety science research politically oriented again.[9] A rich discourse on the politics of risk and safety is needed to supplement questions like "Is it safe?" and "What is the risk?" with questions like "What kind of risks are we *constructing*", "What kind of society do we *want*?" and "How can we get there?". Perrow's warning about inevitable, intolerable risks was silenced far too quickly, and the cultural approach of HRO and the adaptation perspective of RE do not really address the core concern that we may be infrastructuring risk into lifeforms that are extremely hard to reverse – climate change is one example, big-data based autonomous systems[10] another. To make safety science more politically oriented does not hinge on making engineers interested in politics. Indeed, engineers and safety practitioners and scientists exercise politics every day, the question is whose politics they exercise, and how conscious they are about this. A claim in this paper is that they increasingly exercise the politics of the market – which is increasingly outsourced from the parliament. Therefore, the bringing of political orientation back to safety science should be accompanied with a renewed orientation towards societal resilience in the traditional political fora. The shaping of safer societies rests on an intimate relation between science and politics – a relation whose many facets we can witness unfold and develop in contemporary crises like the climate crisis and the corona crisis.

4. Conclusion

Debates in safety science may assume different characters when seen through new theoretical lenses. Debates may take many turns; the parties may manage to partly or wholly convince each other, and adjust their stances accordingly, or the heat of the debate may lead the parties to work even harder developing their arguments, bringing them even further away from each other. The debates I shed light on in this article have not so much taken those turns as they have faded into what we might call an unspoken truce: the NAT/HRO debate went through a deadlock to a dead end[70,71] from which it never really has been revitalised; the HRO/RE debate has never really established an arena or channels where the disagreements can find the nutrients necessary to be really vital,[11] the very few cross-references also reflects this [29]; the Safety I/Safety II debate suffer from the Safety II proponents missing voluntary opponents.

As the present discussion reveals, there lies underexposed elements of controversy underneath the lid of many debates, controversies with a potential to bring reflections on complexity, invisibility and risk – and many more issues, these are only examples that are discussed here – into a more fundamental debate on safety and risk in society. As noted elsewhere [36], there is a lack of engagement in many debates in the safety science community; perhaps we live too well with contradicting views and perspectives at arm's length. By taking epistemology serious and lending support from the wider field of social sciences, we are better equipped to identify scientific controversies, and to advance safety science so that it keeps relevance in rapidly changing societies. What this article suggests, is that this is not done by resolving old debates, but by orienting ourselves towards persistent controversies that resides within them – and indeed also outside them. Three propositions are brought forward here:

- Take complexity seriously, and do not treat it as an unambiguous, independent variable.
- Broaden the perspectives and methodologies for understanding sociotechnical work, by looking to and learning from other research traditions associated with sociology of work.
- Make safety science research politically oriented (again). It is indeed not only a matter of concern for safety researchers how we can live with high-risk system, but also how we can live without them.

CRediT authorship contribution statement

Torgeir Kolstø Haavik: Conceptualization, Methodology, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[9] By being politically *oriented* I mean to be reflexive when formulating research scope and questions, since any research question (and its scope) is entangled with politics. One example, think of how different these two research questions/approaches may be: 1) What are the characteristics of resilient infrastructures? and 2) How can resilience (or risk) be infrastructured? The *research* undertaken in a research project/programme addressing 1) and 2) respectively must be just as objective and epistemologically rock solid, but the framings are very different. And to be politically oriented means to be reflexive on how research framing makes a difference.

[10] Not only technical – also economic, social and political.

[11] This critic has been raised before by [36].

References

[1] Albrechtsen E, Besnard D. Oil and gas, technology and humans: assessing the human factors of technological change. Farnham: Ashgate; 2013.
[2] Ale B. Risk analysis and big data. Safety and Reliability 2016.
[3] Ale BJ, Hartford DN, Slater DH. Variability; Threat or Asset. In: IChemE symposium series no 166. HAZARDS; 2019. p. 29.
[4] Ale BJ, Hartford DN, Slater DH. Variability: for better and for worse in safety assurance. Med Res Arch 2020;8(2).
[5] Almklov PG, Antonsen S. The commoditization of societal safety. J Contingen Crisis Manage18 2010;(3):132–44.
[6] Aven T, Renn O. On risk defined as an event where the outcome is uncertain. J Risk Res 2009;12(1):1–11. https://doi.org/10.1080/13669870802488883.
[7] Ayres RU, Rohatgi PK. Bhopal: lessons for technological decision-makers. Technol Soc 1987;9.
[8] Beck U. Risk society: towards a new modernity. London, UK: Sage; 1992.
[9] Bergström J, van Winsen R, Henriqson E. On the rationale of resilience in the domain of safety: a literature review. Reliability Engineering & System Safety 2015;141:131–41.
[10] Bloor D. The strong programme in the sociology of knowledge. Knowledge and Social Imagery 1976;2:3–23.
[11] Braithwaite J, Wears RL, Hollnagel E. Resilient health care, volume 3: reconciling work-as-imagined and work-as-done. CRC Press; 2016.
[12] Callon M. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. Ed(s). In: Law John, editor. Power, action and belief: a new sociology of knowledge. London: Routledge; 1986. p. 196–223.
[13] Dekker S. Foundations of safety science: a century of understanding accidents and disasters. Routledge; 2019.
[14] Drupsteen L, Guldenmund FW. What is learning? A review of the safety literature to define learning from incidents, accidents and disasters. J Contingen Crisis Manage 2014;22(2):81–96.
[15] Engeström Y, Middleton D. Cognition and communication at work. Cambridge, UK: Cambridge University Press; 1996.
[16] Gusfield JR. The culture of public problems: drinking-driving and the symbolic order. University of Chicago Press; 1981.
[17] Hale AR, Hovden J. Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment. Occupation Injury Risk Prevent Intervent 1998:129–65.
[18] Hilgartner S. The social construction of risk objects: or, how to pry open networks of risk. Ed(s). In: Short James F, Clarke Lee Ben, editors. Organizations, uncertainties, and risk. Boulder: Westview Press; 1992. p. 39–53.
[19] Hollnagel E. Coping with complexity: past, present and future. Cognition, Technology & Work 2012;14(3):199–205.

[20] Hollnagel E. Making Health Care Resilient: from Safety-I to Safety-II. Ed(s). In: Hollnagel Erik, Braithwaite Jeffrey, Wears Robert L, editors. Resilient health care. Farnham: Ashgate; 2013. p. 3–17.
[21] Hollnagel E. Safety-I and safety-II: the past and future of safety management. Farnham, Surrey, England: Ashgate; 2014.
[22] Hollnagel E. Safety-II in practice: developing the resilience potentials. Routledge; 2017.
[23] Hollnagel E, Braithwaite J, Wears R. Resilient health care. London: Ashgate; 2013.
[24] Hollnagel E, Nemeth C, Dekker S. Resilience engineering perspectives: remaining sensitive to the possibility of failure. Aldershot, UK: Ashgate; 2008.
[25] Hollnagel E, Pariès J, Woods DD, Wreathall J. Resilience engineering in practice: a guidebook. Farnham: Ashgate; 2011.
[26] Hollnagel E, Wears R, Braithwaite J. From safety-I to safety-II: a white paper. There silient healthcare net: published simultaneously by the. Australia: University of Southern Denmark, University of Florida, USA, and Macquarie University; 2015.
[27] Hollnagel E, Woods DD, Leveson N. Resilience engineering: concepts and precepts. Aldershot, UK: Ashgate; 2006.
[28] Hopkins A. Lessons from longford: the esso gas plant explosion. CCH Australia limited; 2000.
[29] Hopkins A. Issues in safety science. Saf Sci 2014;67:6–14.
[30] Haavik TK. On Components and Relations in Sociotechnical Systems. J Contingen Crisis Manage 2011;19(2):99–109.
[31] Haavik TK. New tools, old tasks: safety implications of new technologies and work processes for integrated operations in the petroleum industry. Farnham, UK: Ashgate; 2013.
[32] Haavik TK. On the ontology of safety. Saf Sci 2014;67:37–43.
[33] Haavik TK. Sensework. Computer Supported Cooperative Work (CSCW) 2014;23 (3):269–98. https://doi.org/10.1007/s10606-014-9199-9.
[34] Haavik TK. Sensework. Ed(s). In: Le Coze Jean-Christophe, editor. Safety science research: evolution, challenges and new directions. Boca Raton, FL: CRC Press; 2019. p. 103–17.
[35] Haavik, T.K. (Forthcoming). Actor Network Theory and Sensework in safety research. In Kenneth ArnePettersen and CarlMcrae (Ed(s).) Inside hazardous systems - Perspectives on safety and accident research. Boca Raton: CRC Press.
[36] Haavik TK, Antonsen S, Rosness R, Hale A. HRO and RE: a pragmatic perspective. Saf Sci 2019;117:479–89.
[37] Irwin A. Risk and the control of technology: public policies for road traffic safety in Britain and the United States. Manchester University Press; 1985.
[38] Kuhn TS. The structure of scientific revolutions. University of Chicago press; 1962. Original edition.
[39] La Porte T, Perrow C, Rochlin G, Sagan S. Systems, organizations, and the limits of safety: a symposium. J Contingen Crisis Manage 1994;2(4):205–40.
[40] La Porte TR. High Reliability Organizations: unlikely, demanding and at risk. J Contingen Crisis Manage 1996;4(2):60–71.
[41] La Porte TR, Consolini PM. Working in practice but not in theory: theoretical challenges of high reliability organizations. J Public Admin Res Theo 1991;1(1): 19–47.
[42] Latour B. Visualization and Cognition: thinking With Eyes and Hands. Knowledge and Society: studies in the Sociology of. Culture Past and Present 1986;6:1–40.
[43] Latour B. Science in action: how to follow scientists and engineers through society. Milton Keynes, UK: Open University Press; 1987.
[44] Latour B. Circulating References. Sampling the Soil in the Amazon Forest. Ed(s). In: Latour Bruno, editor. Pandora's hope: essays on the reality of science studies. Cambridge, MA: Harvard University Press; 1999. p. 24–79.
[45] Latour B. Is re-modernization occurring-and if so, how to prove it?A commentary on Ulrich Beck. Theory, Culture and Society 2003;20(2):35–48.
[46] Latour B. Reassembling the social: an introduction to actor-network-theory, clarendon lectures in management studies. Oxford, UK: Oxford University Press; 2005.
[47] Latour B, Woolgar S. Laboratory life: the social construction of scientific facts. Beverly Hills, CA, USA: Sage Publications; 1979.
[48] Law J. Ladbroke grove, or how to think about failing systems. Centre for Science Studies, Lancaster University; 2000. accessed 22/06/11. www.comp.lancs.ac.uk/sociology/papers/Law-Ladbroke-Grove-Failing-Systems.pdf.
[49] Le Coze J-C. Are organisations too complex to be integrated in technical risk assessment and current safety auditing? Saf Sci 2005;43(8):613–38.
[50] Le Coze J-C. The foundations of safety science. Saf Sci 2013;67:1–5.
[51] Le Coze J-C. Outlines of a sensitising model for industrial safety assessment. Saf Sci 2013;51(1):187–201.
[52] Le Coze J-C. An essay: societal safety and the global1, 2, 3. Saf Sci 2018;110:23–30.
[53] Le Coze J-C. Safety, model, culture. In (Ed(s).) safety cultures, safety models. Cham: Springer; 2018. p. 81–92.
[54] Le Coze J-C. Safety science research: evolution, challenges and new directions. CRC Press; 2019.
[55] Le Coze J-C. Vive la diversité! High Reliability Organisation (HRO) and Resilience Engineering (RE). Saf Sci 2019;117:469–78.
[56] Nemeth C, Hollnagel E, Dekker S. Resilience engineering perspectives: preparation and restoration. Aldershot, UK: Ashgate; 2009.
[57] Olsen OE, Engen OA. Technological change as a trade-off between social construction and technological paradigms. Technol Soc 2007;29(4):456–68.
[58] Papazoglou I, Aneziris O, Bellamy L, Damen M, Ale B, Manuel H, Oh J. Quantification of risk rates of occupational accidents. Ed(s). Safety and reliability: methodology and applications. CRC Press; 2014. p. 1381–90.
[59] Paté-Cornell ME. Learning from the piper alpha accident: a postmortem analysis of technical and organizational factors. Risk Analysis 1993;13(2):215–32.
[60] Perrow C. Normal accidents: living with high-risk technologies. Princeton, NJ: Princeton University Press; 1984.
[61] Perrow C. Normal accidents: living with high-risk technologies. Princeton, NJ, USA: Princeton University Press; 1999.
[62] Pettersen KA, Schulman PR. Drift, adaptation, resilience and reliability: toward an empirical clarification (special issue on HRO and RE). Saf Sci 2019;117:460–8.
[63] Pinch TJ, Bijker WE. The social construction of facts and artefacts: or how the sociology of science and the sociology of technology might benefit each other. Soc Stud Sci 1984;14(3):399–441.
[64] Rasmussen J. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans Syst Man Cybern 1983;(3): 257–66.
[65] Rasmussen J. Risk management in a dynamic society: a modelling problem. Saf Sci 1997;27(2–3):183–213.
[66] Reason J. The Chernobyl errors. Bull Br Psychol Soc 1987;40(206):1–20.
[67] Reason J. Human error. Cambridge university press; 1990.
[68] Reason J. Managing the risks of organizational accidents. Aldershot: Ashgate; 1997.
[69] Righi AW, Saurin TA, Wachs P. A systematic literature review of resilience engineering: research areas and a research agenda proposal. Reliab Eng Syst Safe 2015;141:142–52.
[70] Rijpma JA. Complexity, Tight-Coupling and Reliability: connecting Normal Accidents Theory and High Reliability Theory. J Contingen Crisis Manage 1997;5 (1):15–23.
[71] Rijpma JA. From Deadlock to Dead End: the Normal Accidents-High Reliability Debate Revisited. J Contingen Crisis Manage 2003;11(1):37–45.
[72] Roberts, K.H. (1993). Introduction. In K.H.Roberts (Ed(s).) New challenges to understanding organizations, pp 1–10. New York.
[73] Rosendahl T, Hepsø V. Integrated operations in the oil and gas industry: sustainability and capability development. Hershey, Pa: IGI Global; 2013.
[74] Skjerve A, Albrechtsen E, Tveiten C. Defined situations of hazard and accident related to integrated operations on the Norwegian continental shelf. SINTEF; 2008.
[75] Stanton NA, Margaryan A, Littlejohn A. Learning from incidents. Safety Science: Special Issue on Learning from Incidents 2017;99(A):1–4.
[76] Star SL, Strauss A. Layers of Silence, Arenas of Voice: the Ecology of Visible and Invisible Work. Computer Supported Cooperative Work 1999;8(1–2):9–30.
[77] Strum SS, Latour B. Redefining the social link: from baboons to humans. Soc Sci Inform 1987;26(4):783–802.
[78] Suchman L. Supporting articulation work (Ed(s).). In: Kling Rob, editor. Computerization and controversy: value conflicts and social choices. San Diego, CA, USA: Academic Press; 1996. p. 407–23.
[79] Sutcliffe KM, Weick KE. Mindful organizing and resilient health care. Ed(s). In: Hollnagel Erik, Braithwaite J, Wears RL, editors. Resilient health care. London: Ashgate; 2013.
[80] Turner BA. The organizational and interorganizational development of disasters. Adm Sci Q 1976;21(3):378–97.
[81] Turner BA. Man-made disasters. London: Wykeham Science Press; 1978.
[82] Turner BA. The sociology of safety. Ed(s). In: Blockley DI, editor. Engineering safety. London: McGraw-Hill; 1992. xvii–475s.
[83] Vaughan D. The challenger launch decision: risky technology, culture, and deviance at NASA. Chicago: University of Chicago Press; 1996.
[84] Vaughan D. The dark side of organizations: mistake, misconduct, and disaster. Annu Rev Sociol 1999;25(1):271–305.
[85] Vaughan D. The role of the organization in the production of techno-scientific knowledge. Soc Stud Sci 1999;29(6):913–43.
[86] Vinnem JE, Aven T, Husebø T, Seljelid J, Tveit OJ. Major hazard risk indicators for monitoring of trends in the Norwegian offshore petroleum sector. Reliab Eng Syst Safe 2006;91(7):778–91.
[87] Wears RL, Hollnagel E, Braithwaite J. Resilient health care, volume 2: there silience of everyday clinical work. Ashgate Publishing, Ltd; 2015.
[88] Wears RL, Roberts KH. Special issue, safety science, high reliability organizations and resilience engineering. Elsevier; 2019.
[89] Weick K. Sensemaking in organizations. Thousand Oaks, CA: Sage; 1995.
[90] Weick KE. Organizational culture as a source of high reliability. Calif Manage Rev 1987;29(2):112–28.
[91] Weick KE, Roberts KH. Collective Mind in Organizations: heedful Interrelating on Flight Decks. Adm Sci Q 1993;38(3):357–81.
[92] Weick KE, Sutcliffe KM. Managing the unexpected: resilient performance in an age of uncertainty. San Francisco, CA, USA: Jossey-Bass; 2007.
[93] Weick KE, Sutcliffe KM, Obstfeld D. Organizing for high reliability: processes of collective mindfulness. Res Organ Behav 1999;21:81–123.
[94] Wiener N. Cybernetics or Control and Communication in the. Animal and the Machine 1961;Vol. 25.
[95] Øien K, Utne I, Tinmannsvik R, Massaiu S. Building safety indicators: part 2–application, practices and results. Saf Sci 2011;49(2):162–71.