Safety Science xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Safety Science

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

Is safety a subject for science?

Erik Hollnagel *

University of Southern Denmark, Odense, Denmark
Region of Southern Denmark, Middelfart, Denmark

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Foundations
Scientific subject
Epiphenomenon
Non-event
Social construct

ABSTRACT

In this paper I will not so much address the status of safety science as a science, but rather address the status or meaning of safety. So instead of entering into a discussion of whether safety science is a proper science – whatever that means – the focus will be on whether the notion of safety itself is a proper subject for scientific investigation or indeed whether safety as such is an appropriate topic or subject for a scientific discipline.

1. Introduction

The invitation, or perhaps the challenge, to submit a paper for this special issue of safety science, questioned the notion of *safety science* as a proper science. Or that was at least how I interpreted the invitation. It was mentioned, for instance, that the status of *safety science* is contested, and this presumably referred to its status as a science. The call also listed several potential controversies, for instance between a 'normative', and a 'descriptive', view, between a 'realist' and a 'constructivist' view, and between viewing safety as 'resulting' or as 'emerging'.

In this paper I will not so much address the status of *safety science* as a science, but rather address the status or meaning of *safety*. So instead of entering into a discussion of whether *safety science* is a proper science – whatever that means – the focus will be on whether *safety* itself is a proper subject for scientific investigation or indeed whether safety as such is an appropriate topic or subject for a scientific discipline.

In order to do so it is necessary to assume that there exists some kind of agreement about the meaning of the term an 'appropriate scientific subject' and therefore also about the meaning of the term 'science'. This agreement need not exist among everyone but must at least be found within a certain community, in this case the community of safety scientists. There is no way of avoiding this thorny issue – short of the irresponsible attitude that takes for granted that we all know what the terms mean and that they mean the same to us all. The issue will nevertheless be given short thrift by resorting to the common definitions that easily can be found in both printed and electronic knowledge repositories. While the Latin word *scientia* means 'knowledge', the modern use of science refers to the ways in which knowledge is pursued, as much as to the knowledge itself. *Safety science* is therefore taken to refer both to what we know about *safety* and to the ways we have built and continue to build this knowledge. In other words, to how we study the subject matter, which in this case is *safety* itself.

If the common definitions are accepted, then a science must have a more or less well-defined topic, focus, or object (phenomenon) that can be studied. It must have a paradigm, as argued by Kuhn (1962). It follows from this definition that astronomy is a science because it studies celestial objects (such as moons, planets, stars, nebulae, and galaxies); that chemistry is a science because it studies the composition, properties and behaviour of matter; that psychology is a science because it studies the mental functions and behaviours of humans; that organisational studies is a science because it examines how organisational structures, processes, and practices shape social relations and influence performance; and at a stretch that even economics can be thought of as a science that studies the production, exchange, distribution, and consumption of goods and services.

According to this way of reasoning, *safety science* is the study of *safety*. But unlike the celestial objects, unlike matter, even unlike mental faculties, organisations, goods and services, *safety* does not represent an agreement on cannot what it is that should be studied, nor can it be said to exist in any concrete or material sense, or to be real (Westenhoff, 2011). Because of this we cannot resolve disputes about what safety is by referring to something that exists independently of our thinking of it, as if it was an object (as the term is used in semiotics). Yet we need to be able to refer to what *safety* is in a way that is open to intersubjective verifiability, we need to have a common agreement on what we should focus on, to avoid falling into the trap of solipsism.

* Address: P.V. Tuxensvej 5, DK-5500 Middelfart, Denmark. Tel.: +45 12345678.
E-mail address: erik.hollnagel@regionsyddanmark.dk

0925-7535/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ssci.2013.07.025

2. The definition of safety

Throughout the ages, the starting point for safety concerns has been the occurrence, potential or actual, of some kind of adverse outcome, whether it has been categorised as a risk, a hazard, a near miss, an incident, or an accident. Heinrich (1929), who by rights must be considered the pioneer of industrial safety, was careful to point out that a distinction should be made between accidents and injuries, where the former denoted the cause and the latter the effect. "There are major and minor injuries, of course, and it may be said that a major accident is one that produces a major injury. However, the accident and the injury are distinct occurrences; one is the result of the other, and in the continued use of the expression 'major accident', and in the acceptance of its definition as one that results seriously, there is a decided handicap to effective work" (Heinrich, 1929, p. 2). Later thinkers have, however, be less scrupulous in their use of the terminology, and the term safety has therefore been used to cover not only the injuries but also the events that lead to them.

Safety is often, indeed nearly always, defined as a condition where nothing goes wrong (injuries, accidents/incidents/near misses) or more cautiously as a condition where the number of things that go wrong is acceptably small. Examples of this definition are easy to find. The International Civil Aviation Organisation, for instance, defines safety as "the state in which harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and risk management" while the U.S. Agency for Healthcare Research and Quality defines safety as the "freedom from accidental injury". More indirect definitions can also be found. As an example, Transport Safety Victoria defines a major incident as "an incident or natural event that poses a serious and immediate risk to safety and includes a derailment of rolling stock, a collision, a fire or explosion". From this one may conclude that if accidents and incidents are a risk to safety, then safety is marked by the absence of accidents and incidents.

Such definitions of safety are, however, indirect rather than direct since safety is defined by what happens when it is absent or missing. Properly speaking, they are therefore definitions of lack of safety (or unsafety) rather than of safety. One consequence of this is that safety management relies on measurements that refer to the absence of safety rather than to the presence of safety. Because the focus is on things that go wrong, there will be something to measure when safety is absent, but paradoxically nothing to measure when safety is present. This has profound practical consequences for how safety is managed, but since that is far beyond the scope of this paper it will not be discussed further here.

The focus on situations where things go wrong, on the absence of safety, is theoretically and scientifically suspect but makes eminent practical sense. First of all because such situations may lead to unintended and unwanted injuries or harm in the form of loss of life and property, disrupted or inefficient performance, etc. Secondly because they usually happen unexpectedly and thereby are a constant reminder of how hard it is to create and maintain the orderly and predictable work environments that we desire so much – for psychological as well as practical reasons. Unexpected and unwanted events such as the collapse of a building or a bridge have been a typical concern in the classical safety thinking. Such concerns have presumably been an integral part of human activity at least since the agrarian revolution around 10–12,000 years ago and has been reinforced many times since. Closer to our time they came to the fore after the second industrial revolution, around 1750. The rapid mechanisation of work in the 19th century led to a growing number of hitherto unknown types of accidents, where the common factor was the breakdown, failure, or malfunctioning of active technology. The mechanisation and industrialisation did not change the nature of the outcomes as such – still a loss of life, material, and property – but it increased the magnitude of the injuries. Hale and Hovden (1998) have characterised this as the age of technology, in which safety concerns focused on guarding machinery, stopping explosions and preventing structures from collapsing. The focus was the risks related to passive technology and structures such as buildings, bridges, and ships. (Petroski, 1992). Seeing technology as the predominant – and mostly also the only – source of both problems and solutions in safety was maintained with reasonable success until 1979, when the accident at the Three Mile Island nuclear power plant (TMI) demonstrated that safeguarding technology was insufficient. The TMI accident forced safety professionals to consider the role of human factors – or even of *the* human factor – and made it necessary to include human failures and malfunctioning as potential risks, first in operation but later also in design, construction, and maintenance (Swain and Guttman, 1983; Dougherty, 1990). In 1986, 7 years later, the loss of the space shuttle Challenger, together with the accident in Chernobyl, made yet another extension necessary. This time it was the influence of the organisation, captured by terms such as organisational failures (Reason, 1997) and safety culture (Guldenmund, 2000).

The history of safety contains several such transitions that occurred when the safety community found itself face to face with accidents that could not easily or comfortably be explained by the existing conceptual framework. In each case, new types of accidents have been accounted for by adding new types of causes (e.g., metal fatigue, 'human error', violations, organisational failure, and safety culture) to the previously existing catalogue. The general concern for safety management has always been to find a cause, or a set of causes, both in order to explain what has happened and in order to propose remedial actions. This way of thinking corresponds to a *causality credo*, which can be formulated as follows: (1) adverse outcomes (accidents, incidents, etc.) happen when something goes wrong; (2) adverse outcomes therefore have causes, which can be found, and (3) treating – and preferably eliminating – the causes will increase safety by preventing future accidents (e.g., Schröder-Hinrichs et al., 2012). An alternative approach would, of course, be to challenge or change the basic underlying assumption of causality, but few have entertained that. We have therefore through centuries become so accustomed to explaining accidents in terms of cause-effect relations – simple or compound – that we no longer notice it. And we cling tenaciously to this tradition, although it has becomes increasingly difficult to reconcile with reality.

2.1. Safety as an epiphenomenon

This way of defining safety indirectly, namely as that which is missing when something goes wrong, sees safety as an epiphenomenon rather than as a phenomenon. (An epiphenomenon is defined as an incidental product of some process, that has no effects of its own.) The primary phenomena are the adverse outcomes and how they come about, and safety is simply a name for the condition that exists when the adverse outcomes do not happen. In relation to the question addressed by this paper, the subject matter of *safety science* is therefore the occurrence – or rather, the non-occurrence – of adverse outcomes (accidents, incidents, and near misses) and their aetiology, but not safety as such. The subject matter is the lack of safety rather than safety. This raises the interesting question of whether it is possible to have a science about something that is not there? In other words, can the object of a science be nothing? (Lest the reader objects, philosophy can study the concept of nothing, but not nothing itself. *Ex nihilo nihil fit.*)

Please cite this article in press as: Hollnagel, E. Is safety a subject for science? Safety Sci. (2013), http://dx.doi.org/10.1016/j.ssci.2013.07.025

It may be countered that safety is not about nothing, but about avoiding adverse outcomes (accidents and incidents), and that safety therefore is the set of methods, principles and practices that have been developed to identify and eliminate (or attenuate) hazards. Yet in that case *safety science* is rather about risks and hazards – meaning the conditions that represent a lack of safety – than about safety.

2.2. Safety as a non-event

The problem alluded to above has been accentuated by the suggestion that safety should be defined as a 'dynamic non-event' (Weick, 2001, p. 335). (Weick actually talked about 'reliability as a dynamic non-event' but the similarity to safety is unmissable.) The meaning of a 'non-event' is, of course, that safety is present when there are no adverse events, i.e., when nothing goes wrong. The meaning of 'dynamic' is that the condition of nothing happening, meaning that nothing goes wrong, cannot be achieved by passive means, by adding layer upon layer of defence and protection, but requires constant attention.

Despite its ingenuity, this definition presents a practical problem, namely that it is impossible to study a non-event. It is impractical but not impossible to count how many times something goes well, and it is almost never done. But is quite impossible to compare two instances of nothing happening – because there is nothing to compare. Weick's definition is nevertheless very useful because it highlights the problems with the conventional understanding of safety. But is not very practical – and was presumably never intended to be so.

The focus on non-events does obviously not mean that nothing happens. Indeed, many things happen, but they succeed rather than fail. This becomes clear if it is rephrased so that safety is defined as 'a dynamic lack of failures'. If we go one step further and replace the 'lack of failures' with 'successes', we arrive at Safety-II as a proper alternative to Safety-I, cf., below.

2.3. Safety as a social construct

Another view, complementary to the above, is that safety (or rather, safe operation) is a social construct (Searle, 1995). This was pointed out by Rochlin (1999), who wrote:

> "The maintenance of safe operation so defined is an interactive, dynamic and communicative act, hence it is particularly vulnerable to disruption or distortion by well-meant but imperfectly informed interventions aimed at eliminating or reducing 'human error' that do not take into account the importance of the processes by which the construction of safe operation is created and maintained" (p. 1549).

One important feature of this definition is the distinction between *safety* and safe operations. While the former may be difficult to define (cf., above) and to some extent intangible, the latter are far easier to talk about and to work with. Safe operations, or operating safely, point to a characteristic way of carrying out the work, hence refer to something that is perfectly observable – it refers to events rather than to non-events. Rochlin did go onto note that it would be a challenge "to identify rules that correlate constructions of operational safety with empirical observations of performance, so as to separate those organisations that construct a representational framework of safety that is accurate from those who only construct the representation, and not the safety" (p. 1558). But the challenge is a concrete rather than a theoretical one. So while it may be impossible for *safety science* to study something that does not exist (i.e., the non-events), it is entirely possible to study a social construct, such as safe operations, even though it is not the kind of study that usually is undertaken. The real difficulty probably is to change the mindset of safety scientists, from a focus on that which goes wrong to a focus on that which goes right.

3. From Safety-I to Safety-II

The understanding of safety that has been put forward here is arguably a straw man, but nevertheless one that is widely accepted and widely practised. The development of resilience engineering, which more or less coincided with the beginning of the 21st century, has, however, led to the formulation of an alternative understanding. In order to distinguish between the two, they have been called Safety-I and Safety-II, respectively (Hollnagel, 2013). Safety-I represents the established understanding that has been described above, which means that safety is defined as a condition where the number of adverse outcomes (accidents/incidents/near misses) is as low as possible.

As technical and socio-technical systems have continued to develop, not least due to the allure of ever more powerful information technology, systems and work environments have gradually become less tractable (Hollnagel, 2010; Perrow, 1984). Since the models and methods of current safety management assume that systems are tractable in the sense that they are well-understood and well-behaved, the available tools are less and less able to deliver the required and coveted 'state of safety'. This inability cannot be overcome by 'stretching' the tools, although this is exactly what happens when they are applied to situations for which they were not intended. Simple linear accident models, represented by Heinrich's domino model, are well-suited to situations that resemble what work was like in the 1920s and 1930s, but not to the 1970s and beyond. Composite linear models, represented by Reason's Swiss cheese model, are well suited to situations that resemble what work was like in the 1970s and 1980s, but not to the 2000s and beyond. Models and methods which require that systems are linear with resultant outcomes cannot and should not be used for non-linear systems where outcomes are emergent rather than resultant. The solution is instead to change the definition of safety so that the focus is on what goes right rather than on what goes wrong. This means that the definition of safety no longer will be 'to avoid or prevent that something goes wrong', or words to that effect, but that it rather will be 'to ensure that everything – or as much as possible – goes right'. Safety-II is consequently defined as the ability to succeed under expected and unexpected conditions alike, so that the number of intended and acceptable outcomes (in other words, everyday activities) is as high as possible. (The astute reader may notice that this is a paraphrase of how resilience engineering defines resilience, cf. Hollnagel et al., 2011).

Following this definition, *safety science* changes from being the study of why things go wrong to become the study of why things go right, which means an understanding of everyday activities. All everyday activities are clearly events rather than non-events, which solves Weick's problem, so to speak. Safety – or more precisely Safety-II – thus becomes an aspect or a characteristic of how systems function, and its presence can be confirmed by looking at well-defined categories of outcomes and by understanding how they came about. The purpose is no longer to avoid that things go wrong, but instead to ensure that things go right. This new understanding of safety explicitly acknowledges that systems are intractable rather than tractable. While the reliability of technology and equipment in such systems may be high, workers and managers frequently trade-off thoroughness for efficiency, the competence of staff may vary and may be inconsistent or incompatible, and effective operating procedures may be scarce. Under

Please cite this article in press as: Hollnagel, E. Is safety a subject for science? Safety Sci. (2013), http://dx.doi.org/10.1016/j.ssci.2013.07.025

these conditions humans are clearly an asset rather than a liability and their ability to adjust what they do to match the conditions is a strength rather than a threat. This shift in focus is yet another argument for the futility of studying 'human error' (Hollnagel and Amalberti, 2001).

A logical, but probably unwelcome, consequence of this way of thinking is that safety no longer is a phenomenon in its own right. Indeed, *safety science* should be about 'safe operation' or 'operating safely' rather than about safety *per se*. The proper subject for scientific investigation is why everyday work succeeds, hence *working safely* rather than *safety*. Working safely encompasses how people are able to adjust what they do to match the conditions of work, how they learn to identify and overcome design flaws and functional glitches, how they learn to recognise the actual demands and adjust their performance accordingly, and how they interpret and apply procedures to match the conditions. The study of this is, however, already done by a number of other scientific disciplines, for instance industrial psychology, social psychology, organisation and management, systems thinking, and resilience engineering. Safety science may be useful as a conceptual umbrella term for what is common to these disciplines, but does not replace any of them. The logical conclusion from the above considerations is therefore that *safety* as it traditionally has been understood (which means Safety-I) is not a subject for science, and that a *safety science* therefore is superfluous. Even if we substitute 'risk' for 'safety', a science of risk would be a pre-paradigmatic rather than a normal science according to Kuhn's definitions. From the Safety-I perspective it will, of course, still be necessary to study how accidents happen and how things can go wrong. But that is the study of accidents, as in accidentology, rather than the study of safety. Conversely, Safety-II studies working safely rather than safety. The object of *safety science* is accordingly how people are able to provide the required performance under expected and unexpected conditions alike.

4. The bottom line

It is possible to summarise the arguments of this paper, by the following simple statements:

(1) When something goes wrong, then there is no safety (safety is missing or not there).
(2) When nothing goes wrong, when things just work as they should, then there is safety.

Therefore:

(3) The scientific study of safety should focus on situations where nothing goes wrong, i.e., where there is safety, rather than on situations where something goes wrong – where there is no safety.

To moderate the previous conclusion, *safety science* should study what is there, rather than what is not there. Safety science should study safe operation or working safely, corresponding to what has been called Safety-II. It should study how people work, individually and collectively, and how organisations function, and do that together with the other sciences that have the same focus but which are based on different principles and have different concerns. Although this would be very different from current practices, it would at least contain the seeds of a proper paradigm, whatever it may be called in the end.

References

Dougherty Jr., E.M., 1990. Human reliability analysis – where shouldst thou turn? Reliability Engineering and System Safety 29, 283–299.
Guldenmund, F.W., 2000. The nature of safety culture: a review of theory and research. Safety Science 34, 215–257.
Hale, A.R., Hovden, J., 1998. Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment. In: Feyer, A.M., Williamson, A. (Eds.), Occupational Injury. Risk Prevention and Intervention. Taylor & Francis, London.
Heinrich, H.W., 1929. The foundation of a major injury. The Travelers Standard 1, 1–10.
Hollnagel, E. (Ed.), 2010. Safer Complex Industrial Environments. CRC Press, Boca Raton, FL.
Hollnagel, E., 2013. A tale of two safeties. Nuclear Safety and Simulation 4 (1), 1–9.
Hollnagel, E., Amalberti, R. 2001. The Emperor's New Clothes, or whatever happened to 'human error'? In: 4th International Workshop on Human Error, Safety and System Development, June 11–12, Linköping, Sweden. Invited keynote.
Hollnagel, E., Paries, J., Woods, D.D., Wreathall, J. (Eds.), 2011. Resilience Engineering In Practice: A Guidebook. Ashgate, Farnham, UK.
Kuhn, T.S., 1962. The Structure of Scientific Revolutions. University of Chicago Press, Chicago.
Perrow, C., 1984. Normal Accidents. Basic Books, New York.
Petroski, H., 1992. To Engineer Is Human. Knopf Doubleday Publishing Group, New York.
Reason, J.T., 1997. Managing the Risks of Organizational Accidents. Ashgate Publishing Limited, Aldershot.
Rochlin, G.I., 1999. Safe operation as a social construct. Ergonomics 42 (11), 1549–1560.
Schröder-Hinrichs, J.U., Hollnagel, E., Baldauf, M., 2012. From Titanic to Costa Concordia—a century of lessons not learned. WMU Journal of Maritime Affairs 11, 151–167.
Searle, J.R., 1995. The Construction of Social Reality. Penguin Books, London.
Swain, A.D., Guttman, H.E., 1983. Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications (NUREG CR-1278). NRC, Washington, DC.
Weick, K.E., 2001. Making Sense of the Organization. Blackwell Publishing, Oxford, UK.
Westenhoff, J., 2011. Reality, a Very Short Introduction. Oxford University Press, Oxford, UK.

Please cite this article in press as: Hollnagel, E. Is safety a subject for science? Safety Sci. (2013), http://dx.doi.org/10.1016/j.ssci.2013.07.025