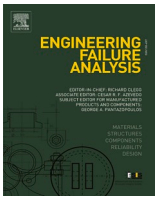
Engineering Failure Analysis 123 (2021) 105315

Contents lists available at ScienceDirect

Engineering Failure Analysis

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



Investigating the implementation of the safety-diagnosability principle to support defence-in-depth in the nuclear industry: A Fukushima Daiichi accident case study

Surbhi Bansal [a], Jon Tømmerås Selvik [a,b,*]

[a] University of Stavanger, P.O. Box 8600, N-4036 Stavanger, Norway
[b] NORCE Norwegian Research Centre, P.O. Box 8046, N-4068 Stavanger, Norway

ARTICLE INFO

Keywords:
Defence-in-depth
Safety diagnosability principle
Usefulness
Nuclear industry
Fukushima Daiichi

ABSTRACT

'Defence in depth' (DID) is a fundamental safety principle applied in several industries, including nuclear. The key is to protect safety critical systems by employing multiple layers of protection, i.e. barriers. The principle states that one single barrier, regardless of how reliable, is insufficient to ensure acceptable safety performance. Obviously then, as the reliability of the layers are associated with the risk of hazardous events, a main safety management activity should be to monitor barrier conditions and performance. However, as experienced in the past, there could be situations where such monitoring is unsatisfactory, challenging the usefulness of the DID. One example, taken from the oil and gas industry, is the 2005 Texas City refinery explosion, where multiple layers of protection failed, resulting in an accident caused by operators with poor situational awareness. Motivated by this assumed weakness, a new principle called the 'Safety diagnosability principle' (SDP) has been suggested for use in the oil and gas industry, in combination with the DID principle. The SDP requires that, for DID to function as intended, any degradation of barriers must be diagnosable and reported. The link to DID makes it also relevant to other industries. In this article, we consider the principle for the nuclear industry. The objective of the article is to clarify the benefits, different ways of implementation, and the potential for using SDP in conjunction with DID in the nuclear industry. To assess the value added, we evaluate the principle against different criteria characterising usefulness. Overall, we find the principle attractive, as the detection and diagnosis of safety–critical events or failures are important for safety management. Having such information strengthens the DID. On the other side, it can also be claimed that acquiring such information is already an implicit part of DID. If so, the SDP adds limited value beyond compliance, i.e. making sure the information is acceptable. We conclude that particularly the relevancy, but also the achievability, related to the use of the SPD, do not point in favour of the principle. A discussion on the 2011 Fukushima Daiichi nuclear accident strengthens our conclusions. The case study indicates that the SDP would not have made the outcome very different. However, as a standalone principle, it might be of greater value. Having reliable information about barrier performance is clearly important to safety management.

* Corresponding author at: University of Stavanger, P.O. Box 8600, N-4036 Stavanger, Norway.
E-mail address: jon.t.selvik@uis.no (J.T. Selvik).

https://doi.org/10.1016/j.engfailanal.2021.105315
Received 3 December 2020; Received in revised form 11 February 2021; Accepted 11 February 2021
Available online 22 February 2021

1. Introduction

Defence in depth (DID) is a safety principle requiring multiple and independent layers of defence, i.e. barriers. Each subsequent layer plays a role in protecting the system, meaning that, always, more than one layer needs to fail for an accident to be possible. It is a principle implemented across several industries. History has also shown, however, that accidents occur, despite systems being designed according to this principle. Saleh et al. [1] examined the Texas City refinery accident and determined that, as a result of misleading information related to barrier conditions and performance, and low situational awareness, operators made the accident possible. Saleh et al. [1] argue that the low awareness originates from the system not being able to provide sufficient information about barrier conditions and the progression of hazardous events. This lack of understanding of what had failed and what was really going on resulted in operators making poor decisions, ultimately leading to the accident. It is acknowledged that, without the availability of updated and reliable barrier information, the value of DID can be questioned. To compensate for this assumed weakness, Saleh et al. [1] suggest pairing DID with a new principle called the 'Safety diagnosability principle' (SDP); see also [2]. The SDP is all about setting up capabilities that reliably detect and report safety-degrading events and barrier failures. It is basically a principle advocating information availability and safety-informed decision-making. For further description, see 2.2 and [1].

The SDP is motivated by the analysis of the 2005 Texas City refinery explosion. The application and conclusions, however, are of a more generic character and linked to the use of DID for various safety management purposes within the oil and gas industry. Saleh et al. [1] also invite other industries where DID is implemented, such as nuclear, to consider the value of implementing the SDP. This suggests that the nuclear industry could face similar challenges regarding the diagnosability of safety barriers. A main objective of this article is to assess why the SDP should also be implemented in the nuclear industry. The key is to assess the usefulness of the principle, which indicates whether it adds value to safety management.

The international standard ISO 12749–5 [3] notes that an objective of DID is to "maintain the effectiveness of the barriers". Clearly, for DID to be effective, either implicitly or explicitly, decision-makers should be informed about safety–critical failures and critical operational aspects related to barrier performance. Otherwise, DID will remain a passive principle, heavily relying on robust barriers. Given that DID encompasses diagnosability requirements, it is possible to manage barriers in a more flexible way, and it will be possible to take actions when and if system reliability is not acceptable. The question is, then, how to achieve such information, as some barriers, for example, could be passive, in the sense that they might have 'hidden failures'. Despite extensive monitoring programmes, some conditions might not be diagnosable before an actual demand. Within maintenance engineering, there is a concept called 'maintenance induced failures' that refers to the possibility that performing, for example, functional testing can cause failures and reduce the reliability. From a system performance perspective, then, collecting reliability information with frequent intervals could be unfavourable for safety, although, if the diagnosability is implicitly already covered, one might question whether there is any need for a second principle on this.

To be clear, we will not give our opinion on the SDP for use in oil and gas and will assume the argumentation and conclusions reached in Saleh et al. [1] to be sound; it is outside our scope to say otherwise. However, it is not obvious that such a principle is needed in the nuclear industry, as it has different sources of hazard (risks), use of technology, operational procedures, etc. [4]. That is where we direct our focus in this article.

As a starting point, we need to define some criteria for what is meant by 'useful' or 'value adding', as a basis for the assessment. For this, we will adopt a set of criteria from Rosencrantz et al. and Sørskår et al. [5,6], used in different contexts to assess the usefulness of other safety principles, i.e. Vision Zero and ALARP (As Low as Reasonably Practicable). These criteria allow us to investigate whether SDP contributes value beyond DID and allows us to capture the relevant pros and cons of the implementation. For specificity, we build the argumentation around the 2011 Fukushima Daiichi nuclear accident. This is one of the most recent events and, with the maximum level 7 on the International Nuclear Event Scale, the most severe nuclear accident since the Chernobyl accident of 1986. In brief, a 9.0-magnitude earthquake off the Japanese coast caused a tsunami that hit the Fukushima nuclear power plant, causing major destruction and the release of radiation to the atmosphere. The plant was designed to withstand waves up to 6 m and was thus unable to stand against the 14-metre-high tsunami wave [7], causing flooding and station-wide blackout at the Fukushima nuclear power plant. In the days following the tsunami, the plant experienced a series of explosions. Several barriers failed. The failure of monitoring and diagnostic instruments impeded the correct diagnosis of the plant and safety system status throughout. We will use this case study to indicate the effect that a hypothetical prior implementation of the SDP could have had for barrier management in this scenario.

The article is structured into six sections. Section 2 outlines the two safety principles in focus: DID and the SDP. This section also clarifies the rationale for using this principle in the oil and gas industry. Section 3 presents and clarifies the criteria adopted for assessing the usefulness of the SDP. Then, in Section 4, we give an overview of what happened at the Fukushima Daiichi accident, the failed safety barriers, and the causal factors. Among these particularly, factors related to the presumed failed diagnosability are identified. In Section 5, we discuss the extent to which improved diagnosability could have prevented the accident or reduced its consequences. Here, the role of failed monitoring systems (e.g. core temperature sensors, water level monitors) is compared with failed mitigatory barriers (e.g. evacuation plans, backup power and water supply) in accelerating the accident. We end the accident discussion by analysing whether restoration of diagnosing capability could have improved the outcome. Finally, Section 6 presents some conclusions and recommendations, based on the identified pros and cons related to use of the SDP in combination with DID in the nuclear industry.

Table 1
Overview of levels in defence in depth [21]

Level 1	Level 2	Level 3	Level 4	Level 5
Prevention of abnormal operation and failures	Control of abnormal operation and detection of failures	Control of accidents within the design basis	Control of severe plant conditions, including prevention of accident	Mitigation of radiological consequences of significant releases of radioactive material

2. Background

2.1. Defence in depth

As described above, DID is the principle of protecting safety or some asset by using multiple layers of successive barriers. The role of the barriers can be visualised with reference to a traditional bow-tie diagram, displaying both preventive and mitigating barriers. It depicts the pathway from causes, through some critical event, to the possible consequences. And it is particularly useful in identifying pathways not following a linear route. The DID complements such a presentation by adding requirements to the barriers displayed or communicated by the bow-tie diagram.

A key when considering DID is that the barriers are independent, and that each layer offers significant protection. It is pointed out for nuclear applications in the fundamental safety principles outlined by the International Atomic Energy Agency (IAEA) [8], "The independent effectiveness of the different levels of defence is a necessary element of defence in depth", meaning that a set of independent barriers must be penetrated for "the asset to be acquired" [9]. It is possible to define DID in different ways, and it has seen some widely discussed developments (see e.g. [9,10]), as might be expected for a principle used for decades in various industries, but the core understanding remains more or less the same.

There are two definitions given in ISO standards, both addressing nuclear applications. ISO 1709 [11] defines DID as "hierarchical deployment of different levels of diverse equipment and procedures (known as barriers) to prevent the escalation of faults to a hazardous condition", which is quite similar to the one given in ISO 12749–5 [3]: "hierarchical deployment of different levels of diverse equipment and procedures to prevent the escalation of anticipated operational occurrences or events". Both standards have adopted and modified the definition given in the IAEA safety glossary [12], where the wording is slightly longer: "A hierarchical deployment of different levels of diverse equipment and procedures to prevent the escalation of anticipated operational occurrences and to maintain the effectiveness of physical barriers placed between a radiation source or radioactive material and workers, members of the public or the environment, in operational states and, for some barriers, in accident conditions."

Typical descriptions of DID comprise terms such as 'successive compensatory measures', 'several layers of protection', 'hierarchical deployment of equipment/procedures', 'depth of penetration', etc. All these terms are associated with the idea of investing in multiple layers aimed at protecting the asset of importance. These are not necessarily safety assets; the asset can be the safety of the workers, society, environment, software, or other hardware (physical) assets. DID is also used for security applications; see [13] for security-related DID definitions. In principle, regardless of application area, the barriers should be effective in managing a system's response to any relevant hazard (human, mechanical and naturally caused events/failures). If one barrier fails to fulfil its intended function, the ongoing hazardous event sequence (e.g. rising reactor core temperature) should be handled in an effective way. The likelihood of severe accidents with serious consequences should be rendered extremely small, with accident prevention being the first priority [14]. For this, safety barriers (such as human, technical or organisational) are employed at every stage (before, during and after) in the event-to-accident escalation path. Barriers at different locations cater for accident prevention, ensuring barrier integrity (or block further escalation) and consequence mitigation [14–16]. However, the principle should be viewed beyond just the barriers, also capturing aspects of control for proper safety management [17], as also stated in the Fukushima Daiichi accident lessons learned [18–19]. The above understanding is summarised in the following three pillars, important for an effective DID strategy [2]:

1. Multiple lines of defence should be placed along potential accident sequences
2. Safety should not rely on a single defensive element (hence the 'depth' qualifier)
3. The successive barriers should be diverse in nature and include technical, operational, and organisational safety barriers (i.e., not only the physical defences).

The three strategy pillars together serve three fundamental safety functions, relevant to the nuclear industry [15]:

- Reactivity control
- Heat removal from the reactor and fuel store
- Confinement of radioactive material

The nuclear industry follows a five-level barrier system, to ensure the above safety functions. This is so that, should one level fail, the subsequent level comes into play [20]. Table 1 gives an overview of these five levels of defence in depth defined by the International Nuclear Safety Advisory Group (INSAG); for notes on the definition of 'defence in depth', we refer to [21].

At the first level, the focus is on typical activities and failures that could have a safety impact. Level 1 refers to main barriers failing,

for example, activating redundant equipment to satisfy a given safety function, or instrumentation giving an alarm when safety-related performance is outside acceptable levels. At the second level, one could have failure of barriers linked to abnormal operational deviations. These are events that do not occur as frequently and might require barriers that have a more passive role in normal operations. The key is to detect and control the situation, so that it does not escalate. At level 3, if a hazardous event occurs, there should be barriers to shut this down in an effective way, to avoid consequences and return to safe operation. Then, for level 4, there should be barriers preventing or inhibiting the consequence development and escalation. Level 5 refers to mitigating barriers related to emergency response, as the final step before the consequences are realised. These levels are discussed in more detail in [21]. The levels can be illustrated by reference to a traditional bow-tie diagram, where levels 1 and 2 are on the left side of the diagram, dealing with causes, the third level being placed around the centre (hazardous event), and levels 4 and 5 being placed on the right side, dealing with mitigating measures and consequences. It is also common to group the levels into three safety layers: hardware, software and management control [19]. Such a combination of barriers, if implemented appropriately, is deemed robust against single or combined failures, unexpected failures and 'beyond design' situations. The key is to ensure independence amongst the barriers. One way to achieve this is by following criteria of diversity, physical separation, and functional isolation [15]. The idea is that independent barriers should not share common causes of failure. It is important that one failed barrier does not increase the probability of other barriers failing. Rather, it should minimise the escalation of deviations during normal operations, particularly to avoid so-called 'cliff-edge effects', i.e. an abruptly large variation in plant condition in response to a small variation in an input [22].

Over the years, the nuclear industry has continuously reviewed the DID content, to ensure it holds as an effective safety principle. This builds on a substantial collective knowledge base that the industry has acquired over the years, including the building, operating and maintaining a variety of nuclear plants, combined with lessons learned from several serious accidents and incidents [22]. The idea of DID has also evolved within different frameworks (such as design-DID, process-DID, and scenario-DID) of nuclear safety; refer to [22] for details. To some extent, this collected experience of lessons learned, observations and use cases contribute to a shared and improved understanding of DID and its value, visible in the regulatory standards of today. Overall, defence in depth is a key concept for better assurance of nuclear safety, by compensating for uncertainties and incompleteness in knowledge [23].

2.2. Safety diagnosability principle

According to Saleh et al. [1], the breakdown of barriers and effects, leading to the 2005 Texas City refinery accident, demonstrates an inherent weakness of DID. It shows that, by adopting this safety principle, one could have multiple independent barriers but still not be well protected. Salah et al. point to the lack of diagnosability, hindering the detection of hazardous states during operation, as a main failure mechanism. Diagnosability refers to the ability to determine whether the system can detect a fault after its occurrence [24]. Poor diagnosability can also be seen as a side effect of redundancy of safety barriers, since it makes the system opaque to the people managing it [25]. For this particular accident, poor system diagnosability left 'blind spots' during operations, concealing the presence of an approaching hazard. This hazard materialised when the conditions in the system exceeded acceptable levels, without the operators being aware of it. The SDP is an initiative to reduce the likelihood of this happening, by requiring an ability to diagnose the hazard build-up concealed by such blind spots.

Saleh et al. [1] outline the SPD as follows: "This principle requires that all safety-degrading events or states that defence in depth is meant to protect against be observable/diagnosable. This principle requires that various features be put in place to observe and monitor for breaches of any safety barrier, and reliably provide this feedback to the operators". See also [26].

The core of the SDP is to reduce uncertainty related to barrier performance, meaning that any barrier should be observable, which in a way gives more control with respect to the issue of uncertainty. The principle requires actions if the conditions are not monitored or observable, given that the information achieved is credible or accurate. It requires reliable information to be available to reflect the barriers' conditions and performance at the relevant time. Facilitating such information allows for actions to make barriers diagnosable or to simply remove them, to avoid a false sense of safety.

A main motivation for this principle is to close the gap between the assumed and actual hazard levels, by increasing awareness of barrier conditions and performance. Its importance for accident prevention lies in the value of the information it supplies and the actions and interventions it spurs [2]. With reference to the Texas City accident, it has been demonstrated that non-compliance with the SDP can degenerate DID into an ineffective defence-blind safety strategy [26]. Violation of the SDP introduces an element of non-transparency regarding barrier effectiveness. Hence, it might lead to a sense of safety by falsely assuming the presence of functional barriers, which can translate into underestimation of hazardous event probabilities. We may end up facing implications of over-confidence in the safety barriers. Factors such as below-expectation barrier performance and a low response time window should obviously be captured by management, to prevent major accidents.

The SDP's usefulness is linked particularly to the left side of the bow tie and the implemented preventive barriers or measures. The availability of these build on the ability to detect and diagnose system conditions. In many situations, this will be necessary for them to perform the required function when needed. For example, there are preventive barriers, dormant in normal operations, such as redundant systems, which might require fault detection as a stimulus to activate them. For manually operated barriers, the sooner the hazardous situation is detected, the quicker barriers can be activated. Further, DID incorporates a need to diagnose safety conditions at different levels (see Table 1). Diagnosability is important to make the operator or decision-maker aware of what is really going on, so that the higher-level barriers are given sufficient attention. Based on this, it could be that the SDP places more weight on preventive compared with mitigatory measures. With robust preventive measures, there is small probability of any mitigating measures being activated in the first place. Based on the analysis of the Texas City refinery accident in [1], it appears that the greater focus is on preventing hazardous events rather than on mitigative measures minimising the consequences. However, that might not be intended.

Table 2
Similarity between rationality and SMART criteria.

Rationality criteria	SMART criteria
Precise	A precise principle is one that is 'directionally, completely and temporally' precise. This corresponds to the 'specific' and 'timely' SMART qualities.
Evaluable	Performance towards the objective stated by the principle should be evaluable. This corresponds to the 'measurability' of progress towards attainment of an objective.
Approachable	Approachability refers to the quality of being 'achievable' or at least approachable to a reasonable degree.
Motivating	Motivating criterion refers to the ability to induce a suitable kind of action by agents. This inherently relates to the 'relevancy' criterion that decides the importance of the objective stated by the principle for business/safety purposes.

The principle should, nevertheless, not be seen as a way of prioritising between preventing (proactive) and mitigating (reactive) measures.

As it is relevant to basically all industries using DID, Saleh et al. [1] also invite the nuclear industry to consider implementing the SPD. The idea is to use this principle to complement DID, but it might also be considered as a standalone principle to strengthen barrier management. However, the SDP has not yet been recommended as a standalone principle, i.e. for situations where DID does not apply. In this article, our focus is on using the principles in combination, meaning that the usefulness or added value of the SDP comes from ensuring informed use of DID. Implicitly, it means that safety decision-making could be improved and could lead to different outcomes, compared with situations with no reference to the SDP.

The need for the SDP is motivated by past events and experience using DID in the oil and gas industry. This is an industry where barrier management overall is given a high level of attention, and where it is recognised as important to observe barrier conditions and performance, and update barrier reliability estimates, to demonstrate that performance satisfies the required safety integrity levels. Especially, there is much focus on barriers in systems with major accident potential. Despite this, for example as regards well design, safety–critical equipment could be installed downhole with limited or complex monitoring options. The oil and gas industry monitors several hazards due to the complex nature of operations that require constant vigilance. There is a wide spread of production activities taking place at several distinct locations, and implementing the latest technology to increase profitability is a common practice [4]. Hydrocarbons need to be moved across units (for example, from offshore platform to gas extraction unit to refinery), and their control is usually more decentralised compared with operations in the nuclear industry, where there is perhaps also less variability in the type of operations, while the potential worst-case consequences of accidents are considered less likely and more severe. Nuclear power plant operators typically have a greater time window to respond during disturbed conditions [4]. There are differences, obviously, but there is nothing in the operational differences to suggest that the SDP should not be transferrable from oil and gas to nuclear.

3. Usefulness assessment criteria for SDP

In the nuclear industry, DID has a role guiding managerial decisions about the sufficiency of levels of protection against the radiation risk. The idea is that the SDP complements the DID, by ensuring a higher focus on quality information feedback related to barrier performance. As a main safety objective is to have functioning barriers at any time, such information is seen as important for barrier management, meaning that, clearly, there are positive aspects. But we should also consider arguments for not implementing the SDP, which will contribute to a more nuanced evaluation of the principle, covering both pros and cons. For example, depending on the system considered, it might be challenging to achieve diagnosability in practice; see e.g. [27].

To assess the overall value of the SDP as a key principle for nuclear applications, we need an appropriate instrument: one that allows us to systematically evaluate its usefulness. What we look for is a set of criteria that can be used to assess whether the quality and value of the information provided by implementing the SDP are sufficiently in favour of the principle, in other words: how the principle influences safety management quality.

For a suitable set of criteria, we refer to Sørskår et al. [6], who use a set of criteria adopted from Edvardsson and Hanson [28] to assess the appropriateness of combining two other key safety or risk management principles, i.e. the ALARP and the Vision Zero principles. For the assessment, four rationality criteria (i.e. precision, evaluability, approachability, and motivating) are used to evaluate relevant aspects. The criteria allow for a consistent and transparent evaluation, while covering the main aspects of risk and safety management.

The four criteria suggested in [6] capture basically the same aspects as the criteria given by the SMART acronym: specific (S), measurable (M), achievable (A), relevant (R) and timely (T), enlisting them as an alternative set of criteria for appropriate quality [29–30]. Although the SMART criteria in [30] are not demonstrated specifically for safety principles, we interpret the two as interchangeable. Table 2 shows the correspondence among their criteria (refer to [6,29–30] for more details).

As can be seen from Table 2, the two alternatives prescribe similar criteria. This means that there are no practical implications of using one set over the other. One should arrive at the same conclusions, irrespective of which set was adopted for the assessment. The SMART framework is clearly the one most cited among the two and considered the most recognised. It is intuitive and quite simple to use in practice, and we will adopt it for our assessment of the SDP in this paper.

The five SMART criteria are further clarified below:

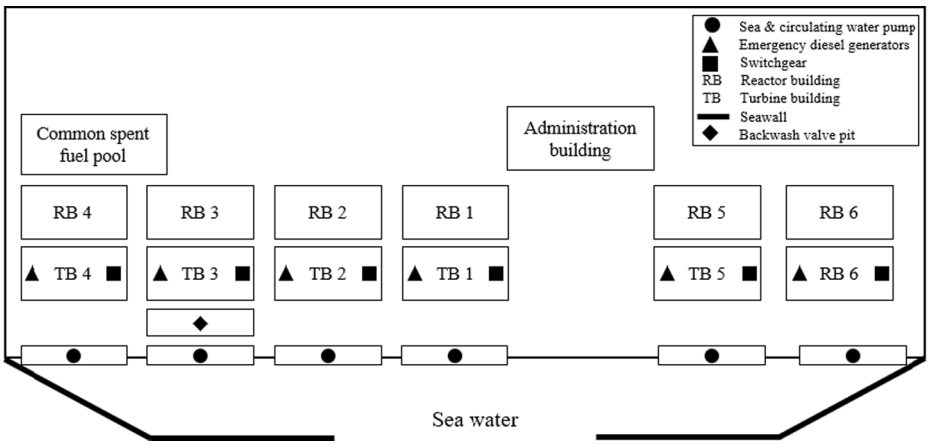


Fig. 1. Fukushima Daiichi nuclear power plant layout.

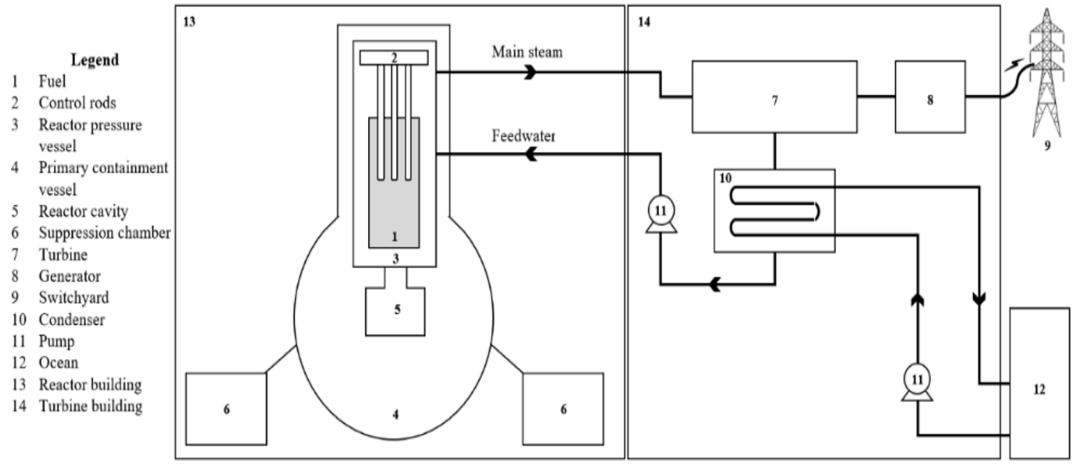


Fig. 2. Setup of the NPP unit.

- *Specific:* The objective of the principle should be precisely and clearly defined. The implementing agents must have a clear understanding, to be able to use it consistently.
- *Measurable*: It should be possible to rationally measure the progress towards or achievement of the objective. Whether the objective is met, where we currently stand, and if we are going in the right direction, should be evaluable.
- *Achievable*: This refers to the degree to which the principle/objective is practically achievable. It concerns factors such as cost, knowledge and practical limitations affecting the certainty of achievement.
- *Relevant*: It should contribute to the organisation in a meaningful way, i.e. add value. The significance will be affected by conflicts or overlap with other business objectives and goals. A relevant principle will also motivate the agents to work for it persistently.
- *Timely*: The principle should have a time horizon in which the objective should be achieved.

The SDP should satisfy all these criteria to prove its informational value to DID and to demonstrate added value for overall safety management. This will serve as an input to evaluate its suitability, in combination with DID, for the nuclear industry.

4. Presentation of case for the SDP assessment

For more specificity, we will refer to an actual accident case scenario as a basis for the discussion. Several of the aspects related to the SMART criteria make little sense without such a reference, particularly achievability and relevancy. Without a more practical context for the discussion, it is difficult to conclude on its actual usefulness. Thus, before moving into an assessment of the SDP using the SMART criteria, we introduce a case based on the 2011 Fukushima Daiichi nuclear accident.

The Fukushima Daiichi plant and process design were guided by the DID principle. However, as history shows, DID's implementation could not prevent the accident from materialising. Below, we investigate whether the SDP would have made a significant difference to the accident outcome. We will use the findings from what happened in the discussion (in Section 5), along with arguments

Table 3
Safety barriers at Fukushima Daiichi units.

Safety barrier against uncontrolled reactivity
Control rods – Scram system to shut down reactor
Safety barriers against reactor heating during operation
Condenser – Cools the feedwater that keeps fuel rods covered
Fuel pool cooling – Spent fuel (in the storage) kept submerged in water
Safety barriers against containment breach
Fuel protection – Zirconium cladding to protect fuel against corrosion
Primary containment vessel – Houses the RPV with nuclear fuel (primary containment barrier)
Reactor building – This concrete building serves as the secondary containment barrier between PCV and external environment
Safety barriers against loss of coolant event
Reactor core cooling – Sprays cooling water on top of the reactor, high-pressure injection system
PCV cooling – Sprays cooling water inside the PCV
Coolant cooling – Isolation condenser, Residual heat removal system, Suppression chamber
Safety barriers for other hazards
Hydrogen release – Hydrogen detection and removal system in the RPV
Fire hazard – Fire protection system (also a backup system for core cooling under accidents)

that can be given on an overall basis for the nuclear industry regarding the SDP. The discussion on its value-adding potential linked to this accident depends to some extent on the findings, but we might not necessarily be able to draw generalised conclusions based on this one accident scenario alone. However, should we conclude that the principle lacks usefulness for this scenario, there will be strong reasons to question the rationale for giving it a key role in the safety management of other nuclear power plants.

4.1. Fukushima Daiichi nuclear power plant overview

Fukushima Daiichi nuclear power plant (NPP) is located on the eastern coast of Japan. Fig. 1 depicts the plant's layout. It has a total of six units (1–6). Units 1–4 are located on the left and the rest are on the right. Each unit has a reactor building (RB), a turbine building (TB), an emergency diesel generator (EDG) and relevant switchgear. The units share a common spent fuel building to store a large amount of fuel assemblies. The pumps located in front are used for pumping sea water and circulating water in the units. The administration building and the emergency response centre are in a seismically isolated building, located behind the units at an elevation. A back-wash valve pit used for filtering water is located in front of unit 3. The site has a seawall, to protect against tsunami waves of a height of up to 5.5 m. It opens directly onto the ocean.

The setup of a typical unit in the power plant is shown in Fig. 2. The unit has two sides: a reactor building and a turbine building. The two sides together run a closed-loop steam cycle. The cycle begins with a nuclear fission reaction inside the reactor pressure vessel (RPV). The RPV is housed in the primary containment vessel (PCV) on the reactor side. The radioactive fuel in the RPV absorbs neutrons, triggering a chain reaction that releases energy. The process reactivity is controlled by control rods and immersing the fuel in water. The PCV is connected to the suppression chambers that store water to manage the reactor pressure. The nuclear reaction generates energy in the form of heat. The RPV has incoming water through a feedwater line. The generated heat vaporises this water, and it travels through the main steam line towards the TB. Here, the steam drives the turbine, so that a generator can produce electricity. After driving the turbine blades, the steam is condensed into water by a condenser. The condenser uses pumped ocean water as its cooling medium. The water is recirculated to the reactor side via the feedwater line, and the cycle keeps repeating. Clearly, ensuring a consistent water supply is important, as it plays multiple roles as a working fluid, coolant and moderator of reactivity.

4.2. Overview of safety barriers

The Fukushima Daiichi plant employed the defence-in-depth principle as its fundamental safety principle. It had three main barrier levels, as against the five levels prescribed in the IAEA standards. The plant should operate safely during normal circumstances, as well as under emergency conditions. For this, several barriers for core cooling and radioactivity containment were ensured. Table 3 lists the safety barriers and their corresponding functions below:

4.3. Accident sequence

Explosions at the Fukushima NPP spanned several days, following a complex sequence of events. The plant supervisors, operators and government authorities were unable to gather information about these events in time. We now look at the accident sequence that led to the explosions.

4.3.1. 4.3.1 initiating event sequence

An earthquake of 9.0 magnitude took place on 11 March 2011, off the Pacific coast of the north-eastern Japanese mainland [31]. The epicentre was 24 km deep into the Pacific Ocean and 180 km from the Fukushima Daiichi NPP [32]. On the incident day, units 1–3 were operational, while units 4–6 were in different stages of planned maintenance.

Unit 4: fuel offloaded to spent fuel pool and emitting a large amount of decay heat

Units 5 and 6: fuel assembly inside the reactor core but emitting low decay heat

The two-minute-long earthquake damaged the power transmission and distribution systems across the region. Fukushima NPP experienced a power outage. The power interruption triggered the automatic emergency response system and stopped the nuclear reaction in units 1–3. Their nuclear cores kept emitting decay heat in their surroundings, raising the temperature and pressure. For a safe halting of operations, a cold shutdown had to be achieved. Cold shutdown is the stage at which, after a few hours of reactor shutdown, actively cooling with recirculated water drops the temperature below 100 °C, such that active cooling is no longer needed, and the reactor becomes passively safe [33].

The earthquake triggered a loss-of-offsite-power (LOOP) event in the plant. This refers to the loss of AC power at the plant. LOOP automatically initiated the onsite EDGs to supply the necessary AC power to the units (1–3). Consequently, the units could begin using the isolation condensers to cool their reactor cores. Their temperature and pressure started lowering immediately. The earthquake also triggered the tsunami waves. Shortly after restoring the emergency power, the plant was flooded by tsunami waves of 16 m height. The 5.5 m seawall was entirely ineffective in preventing site inundation. The flood water entered the reactor, turbine and service buildings. Equipment necessary for ensuring the cooling function, such as pumps, EDGs, motors, power connections, switchgear, etc., were either damaged or immersed in water. The NPP had now also lost its emergency AC power source. This caused a station blackout, a specific event where the plant units experience a loss of AC power for more than five minutes [12]. The offsite emergency response centre and Japanese ministry declared a nuclear emergency.

4.3.2. Consequence sequence

Units 1, 3 and 4 shared a common sequence of events leading to explosions in their respective reactor buildings. These explosions spread out over several days following the tsunami. Given the similarities among their accidental path, we limit our analysis to unit 1, which was the first unit to experience an explosion.

The earthquake had caused the LOOP event. This triggered several emergency response systems: (1) The loss of AC power automatically started the emergency diesel generators. (2) The ventilation system stopped working, and the temperature and pressure inside the containment vessel started rising. The operators diagnosed this and started the cooling system manually. (3) After being shut down, the reactor became isolated from the turbine building's condenser cooling system; its rising pressure automatically started the isolation condenser (IC) system. The IC started removing the residual heat from the PCV. After some time, the IC was manually stopped, as it was decreasing reactor pressure and coolant temperature too rapidly. The NPP's safety barriers were operational, diagnosable, and the situation was now under control.

However, the earthquake was shortly followed by several tsunami waves. The tsunami flooded the basement of the reactor building. The emergency generators, DC panels and battery units located there were inundated. Unit 1 lost both the onsite AC and DC backup power. AC power was crucial to run the safety barrier equipment; the DC power supply was vital for plant safety, as it was needed for instrumentation and control and supplied AC power from inverters to a small number of essential components [32]. The tsunami had the following consequences:

(1) *Loss of backup AC power:* resulted in lost emergency core cooling barriers.
(2) *Loss of backup DC power:* Operators lost instrumentation, alarms and sensors that monitored the reactor water level, reactor pressure, cooling barriers' status, temperature and water level in the spent fuel pool, and status of the IC system.

The reactor lost all the cooling systems and the power necessary to energise and monitor them. Without the cooling function, the containment started to be pressurised by the evaporating water. As the water level dropped, the core would soon become uncovered. The heated core, if unchecked, could melt down and risk radioactive release.

Dissipating the decay heat became a priority in unit 1. The decay heat could accelerate water evaporation and reduce the water level in the core. If this evaporation remained unchecked, the nuclear core would be uncovered, overheated, and might end up in a core meltdown. Loss of AC/DC power due to a blackout triggered a downward spiral of events. The operators could not ensure the core cooling function, as it ran on electricity. They faced a twofold challenge. Firstly, the critical pumps and valves to achieve cold shutdown could not be operated, due to a loss of AC power. Secondly, there was uncertainty about the reactor status, as the unavailability of DC power rendered the instrumentation useless. They decided to initiate their efforts to first arrange power to run the equipment.

They started formulating strategies for barriers that could stop the potential nuclear fuel degradation. For a short duration, the reactor water monitor activated and displayed a decreasing water level in the RPV. So, the team decided to cool the core by injecting water. They started arranging alternative equipment (such as the fire protection system, fire engines and freshwater tanks) for this, given that the existing cooling barriers had been rendered powerless. Additionally, there were repeated attempts to start the IC. The IC system condensed the incoming reactor steam pipeline by submerging it in a cold-water tank. As mentioned above, this system had been shut down just before the tsunami arrived. However, loss of AC power post-tsunami meant that its availability was unknown. The operators tried to restart it, believing that the valves inside the containment that routed steam to the IC were open. This assumption turned out to be wrong, when the IC failed to start. The timing and sequence of power loss had unknowingly closed the valves.

Fearing a degradation of the core, the operators had to manually read the reactor pressure, by visiting the reactor building. They confirmed that core pressure was increasing. By this time, the alternative water injection arrangement was complete, but it could not be initiated. High core pressure conditions rendered the alternative low-pressure water injection impossible. In the meantime, temporary batteries were used to restore DC power and energise the indicators. The readings on the water level monitor indicated that the reactor core was submerged. However, investigation reports suggest that the level indicators were unreliable [13].

After some time, two operators detected radiation outside unit 1, using their personal dosimeters. This was a sign that the core had started degrading, possibly due to low water level. As the radiation started spreading to the main control room, the failure of containment barriers also became a likely scenario. By the end of day one of the accident, the drywell pressure (inside the reactor) was found to be exceeding its maximum design pressure. This high pressure was a warning of an exceedingly critical situation in the unit. The site superintendent decided to vent the PCV to reduce this abnormal pressure level. This was also necessary to resume water injection. They communicated this to the Japanese government, who allowed the venting after residents in a 3-km radius were evacuated. Even after evacuation was complete, the ventilation kept on being delayed.

On 12 March, the following day, the operators managed to start water injection at 0400 h, using a fire truck, which fetched water repeatedly from a freshwater tank. In the following hours, the operators noticed a drop in the containment vessel pressure, without any established ventilation paths. This observation, coupled with a significant increase in radiation dose rate, suggested that the primary containment was failing. In response to this, the government extended the evacuation zone to 10 km.

After a few hours, the workers were able to establish a continuous water injection line between the freshwater tank and the reactor. Although the team had clearance for manually venting the PCV, the ventilation had still not begun. Either the operators were forced to abandon the reactor building as a safeguard against radiation exposure and recurring tsunami threat or they faced challenges in opening the valves manually. After a few hours, they finally managed to open the PCV vent line valves. The pressure venting was done successfully, as a reduction in PCV pressure was observed. By 1530 h, AC power restoration, water provisions and core cooling supplies had been re-established in the unit. However, before they could be used, there was an explosion in the unit 1 reactor building topside. The explosion did not, however, affect the PCV. The source of the explosion is attributed to a hydrogen-air reaction. A reaction between zirconium (nuclear fuel cladding) and water under high temperature had released hydrogen gas, which had, unbeknownst to anyone, escaped to the reactor building via some unobserved path. There, it mixed with the air, causing a violent explosion. Being exothermic in nature, the hydrogen gas reaction produced heat that further accelerated fuel heating [8]. This released more radiation due to core melting, in addition to the radioactive gases released by the explosion. The explosion's pressure damaged the power cables and injection lines laid down for units 2 and 3. In the following days, unit 3 had a hydrogen explosion on the top floor of its reactor building. This was followed by another explosion in unit 4, wherein hydrogen had leaked through a vent from unit 3. Unit 2 did not experience an explosion, despite a damaged reactor core and pressure build-up. The investigators believe the opening of the top floor blow-out panels, due to the explosion in unit 1 nearby, and the lower hydrogen gas generation, to be the possible reasons [8]. The ceiling holes were also potential venting outlets for hydrogen gas accumulating inside the structure.

For further details, we refer to e.g. [19,34].

5. Discussion - assessment of usefulness

5.1. Basis for the discussion

In this section, we will use the above presented case to discuss arguments for and against complementing DID with the SDP principle. We will use the SMART criteria (see Section 3) as a basis for the discussion. The discussion will draw on the experiences from the Fukushima Daiichi accident. This will provide insights into the potential role of the SDP in nuclear accident situations.

5.2. Specificity discussion

This criterion can be assessed on a general basis for nuclear applications and is not specific to the scenario above.

According to definition, the SDP requires that all safety degrading events or states that DID is meant to protect against be observable or diagnosable. In other words, the principle requires the implementation of observing or monitoring features that look out for safety barrier breaches and reliably provide feedback to the operator. The SDP's precision lies in the clarity of its objective and direction to the implementing agent, by requiring actions if this is not fulfilled.

The principle allows for two ways of interpreting the objective: moderately and strictly, of which the moderate objective is substantially less demanding and requires that barrier degrading events are diagnosed and reported through feedback. For consistent implementation, monitoring features should be set up. The features should reliably provide information whenever DID-relevant events cause a safety barrier breach.

The stricter version of the objective leans towards a more extreme safety perspective. It requires the system to monitor the complete state of barriers. This implies that all the status parameters of a safety barrier need to be observable, not just the information about its breach event. Then, the combined scope of monitorable events and states increases exponentially. The rationale is that the barriers with even marginal deviations from the normal operating conditions may lead to a potential barrier breach. The operator should have the maximum amount of information to predict a barrier failure considerably in advance. This will ensure the availability of a longer response window to the operator. The choice of moderate versus stricter SDP objective will depend on considerations such as risk appetite, cost-benefit evaluation, budget constraints, technology challenges, etc. This requires a managerial review and judgement and has been left to the management, as the principle cannot guide on this aspect.

Overall, the principle is seen as sufficiently specific, with a flexibly defined objective. It also provides a definite direction for the actions to achieve the objective. We argue that the principle is sufficiently specific.

5.3. Measurability discussion

Measurability is mostly a matter of which information it is technically possible to collect regarding barrier performance. Although it somewhat depends on the type of barrier, we will be able to draw inferences here based on general barrier understanding. Basically, what we want to know is whether there are obstacles hindering us in monitoring or collecting information on barrier conditions.

The level of barrier diagnosability should be measurable through some metric. To achieve this, we require information such as how many barriers are currently monitored and, amongst these, the number of states or critical events, or the development of degrading processes. But collecting such data can obviously be challenging. The size of the state space would increase exponentially with the system's complexity [35], especially for the stricter SDP objective. The analyst evaluating this metric might have difficulties in comparing the captured state space versus the real state space. Further, all these events/states need to be simulated to count the diagnosable fraction, which is quite challenging. This raises uncertainty about the background knowledge supporting this metric. It can be claimed that any measurement or evaluation made without the knowledge of this uncertainty would be meaningless. Instead, feedback or knowledge of past results can help in measuring and improving performance towards the objective [30]. Trend indicators can measure this progress. For example, for a nuclear reactor with a history of hydrogen gas leaks, an increasing trend of undiagnosed or delayed detections indicates poor diagnosability. The management implementing the SDP can then use this indicator to take actions that improve the diagnosability level in the future (e.g. installing gas detectors at the barriers and hidden escape paths). Such trend indicators also require careful judgement, especially when compiling and evaluating trends for normal operative periods or zero-missed detections.

Trend indicators could be useful in quantifying and assessing the system's ability to observe specific failures and events. Besides, the monitoring ability can be claimed to be simply a matter of cost and not really an issue with respect to the measurability. Overall, this ensures that the SDP's objective is measurable, and we conclude that the measurability criterion is satisfied.

5.4. Achievability discussion

The achievability criterion is highly scenario-specific. In a way, this criterion addresses the core of the principle: whether it is practically possible to obtain the barrier information with high confidence. It is a matter of removing uncertainty related to the barrier performance, while also considering the available resources and other business objectives.

Safety barriers experiencing failures are particularly important for this discussion. Motivated by the case presentation in Section 4, we focus on the performance of the following three barriers:

- The reactor core cooling barrier
- The containment integrity and hydrogen removal system
- The human-organisational barrier

A failure of these barriers was significant for the accident. For each of these barrier failures, we first consider the barrier monitoring capability already present (without following the SDP) and why it failed. Then, we will consider the potential benefits of the SDP: whether its diagnosis information had the potential, retrospectively, to avert the accident.

5.4.1. Loss of reactor core cooling barrier

Right from the start of the accident, the plant lost its normal and emergency core cooling barrier systems. The plant units were equipped with several sensors and instruments to monitor their status. Water level and temperature monitors were used to observe the barrier effectiveness against the accumulation of process decay heat. Additionally, valve status (open or closed) and activation indicators provided information on the barrier cooling's availability or failure. The units ensured diagnosability to a large degree, without mandating the SDP in the first place. This came from the diagnosis and monitoring requirements of DID. Following the tsunami-induced power blackout and site inundation, most of the units lost their safety barriers beyond defence level 2 (see Table 1 for description of levels).

In retrospect, let us consider that the SDP was applied, such that all the monitoring features were functional. Often, normally reliable instrumentation becomes untrustworthy under extreme operating conditions of high pressure, temperature, radiation, etc. Then the reliability of the diagnosis received during accidental situations becomes uncertain. This also happened in the Fukushima accident. The erratic monitoring instrument readings misled the operators. Unit 1′s water level indicator was key to monitoring and confirming the core cooling barrier's status. The instrument's unreliability became known only after the operators discovered that the actual reactor conditions and the displayed readings were incompatible. This uncertainty caused a loss of response time and induced stress among the operators. They made poor decisions that later required additional resources to retract. The operators had to physically verify the reliability of the indicators and lost precious time. They eventually shifted priorities towards re-establishing the integrity of safety barriers and arranging external help. The likelihood of unreliable diagnosis, which deteriorates further as the operating conditions become adverse, undermines the usefulness of the SDP.

5.4.2. Loss of containment integrity and hydrogen removal system

The hydrogen gas leaked from the PCV following unknown paths in unit 1. As per DID's monitoring requirements, the units were equipped with hydrogen detection instruments. These monitored the hydrogen level in the PCV that was filled with inert nitrogen gas as a barrier against explosion. But the plant's DID barriers were not designed to prevent hydrogen gas migration from the PCV to the

reactor building. This was due to the assumption that hydrogen could not leak out of the PCV, which was the only standing barrier preventing hydrogen gas from leaking outside. However, eventually, the combination of core damage, high containment pressure and temperature compromised the containment, allowing hydrogen to escape from the PCV [13]. It is estimated that gaskets, flanges, cableways etc., weakened by high temperature, were possible escape routes that breached the PCV's leak seal and integrity [8]. As a result of this seriously flawed assumption, hydrogen gas build-up in the unit 1 reactor building remained hidden, as there were no monitors to detect it.

The hydrogen level monitors inside the PCV were unavailable due to power outage. The RPV could have accumulated 10,000 m^3 hydrogen in just half a day, due to the high decay heat soon after the reactor tripped [36]. The management blindly relied on the inert atmosphere and ventilation to prevent hydrogen accumulation and leakage. The operators focussed their efforts on core cooling and pressure venting rather than safely disposing of the hydrogen gas. Even the emergency response procedures did not emphasise hydrogen monitoring outside the PCV, despite it being a possibility. We know that the SDP requires that DID-relevant events should be diagnosable. Then, even if the SDP were implemented retrospectively, the units would not have had features installed to observe the PCV-barrier breach. The hydrogen gas breach was not anticipated in the DID barrier design. We can infer that there is a possibility that certain safety-degrading events/states are not within DID's scope. The SDP should have a broader scope, addressing such unaccounted-for hazardous events and unjustified assumptions. Then it could add safety-relevant information that is truly complementary to DID.

Even if the hydrogen detectors were functional, it is possible that hydrogen gas remained undetected. The PCV has a large complex surface area with several leaking paths. Unknown to anyone, gases may accumulate in hidden pockets and pipes for a long time. There is uncertainty about the diagnosis, as it would depend on the location of diagnosing instruments, their range and operating limits. Additionally, while the global containment pressure may remain below a certain safety level, a higher local concentration sensitive to hydrogen distribution may damage specific containment components, internal walls, and safety equipment [37]. This also affects the reliability and timely availability of the diagnosis. An improper design or poor positioning of diagnosing features can affect the extent to which the SDP can be successfully implemented.

After unit 1′s explosion, the operators feared hydrogen explosions in other units. Even in the absence of diagnosability, they logically concluded that hydrogen containment barriers had failed in unit 3, which later turned out to be correct. The likelihood of a high hydrogen level concentration causing an explosion was predicted to be high. The operators were helpless and could not act on this information. The plant personnel did not have access to control equipment, and hydrogen gas ventilation was delayed. The presence of radiation, lack of light source and risk of hydrogen ignition prevented ventilation. Operators were waiting for the arrival of special equipment for cutting holes in the roof and knocking out the panels. Before it arrived, unit 3′s building top had exploded. In such a situation, even if the hydrogen state and its barrier failure had been diagnosed due to SDP compliance, it would not have prevented the explosion from happening, due to ill preparedness. Instead, the timely availability of mitigatory measures, to stop the event escalating, would have had a positive effect.

Unit 4 had an unexpected hydrogen explosion, even though it was not operational to produce hydrogen gas. It received hydrogen gas from a reverse flow from unit 3, via the piping arrangement connected to a common vent stack. One design feature which may have prevented or mitigated the migration of hydrogen is backflow dampers, which were not included in the unit 4 venting system design [13,38]. This is among those scenarios where a robust barrier design, rather than its failure diagnosis, needs to be emphasised. This does not undermine the need to monitor critical barrier states, but we need to compare the SDP's usefulness with mitigatory measures' effectiveness against such hidden hazardous event escalations.

5.4.3. Failure of human-organisational barrier

In Fukushima's case, the failure of the human-organisational barrier and the safety culture played a critical role in the failure of DID. The management of Tokyo Electric Power Company (TEPCO), the nuclear power plant's operating company, did not adopt a strict accident management strategy which could have prevented the simultaneous lack of power availability in all units [13]. Their managers also lacked experience and did not consider the importance of updated risk knowledge. Before the accident, a study had already revealed the likelihood of experiencing a tsunami beyond the Fukushima's handling ability. The organisation ignored the implications of such a study, even though the plant was under-designed. TECPO never addressed the possibility of a prolonged, total loss of power, which led to unpreparedness [38]. The poor safety culture is also visible in the continued use of outdated reactor design, improper placement of emergency generators, compact plant design to reduce land cost, other relaxed safety features, etc.

The Japanese government and regulatory barriers had also weakened. The regulators lacked the power to enforce new requirements emanating from operating experience in other parts of the world. The government had no provisions to manage an extended and widespread loss of power, since they assumed that the power transmission lines would go online quickly. These barriers' failures are difficult to detect and DID does not address them. The failure of these invisible non-technical barriers has more devastating consequences for accident escalation. The SDP lacks guidance on how to monitor the organisational barrier failures; see [26]. It does not add any value in diagnosing these barriers' failure.

5.4.4. Achievement of diagnosability

The SDP's objective is that the implementing agents should develop a system that diagnoses all the safety barrier breaches and delivers this information reliably to the operator. To assess the achievability of this principle, we need to address the uncertainties associated with diagnosability. These uncertainties may arise due to physical limitations, systemic risks, invalid design assumptions, and poor background knowledge. Thy can severely limit the ability to achieve the objective. In other words, targeted actions may have a less than desired effect on the progress towards the objective. For the SDP to satisfy these criteria, we need to evaluate whether diagnosability is actually achievable.

One of the important aspects for achieving the SDP's objective is the reliability of diagnosis feedback. Reliability is associated with multiple aspects such as timeliness, durability, accuracy, precision, etc. In Fukushima's case, negative externality and organisational factors led to a prolonged power interruption. This power blackout was a common cause failure event for the safety barriers and their monitoring instruments. Even though their instruments were reliable, accuracy-wise, they became unavailable and ineffective during hazardous conditions. Likewise, in risky and complex systems, the diagnosing features can simultaneously fail, along with the safety barriers, due to a common failure event (such as a tsunami, in the case of the Fukushima accident). Then, compliance with the SDP may not improve the situational awareness, as it claims. Safety diagnosability, even in the presence of reliable monitoring features, can, in some situations, be difficult to achieve.

In Fukushima's case, we saw that the failure of the reactor cooling barrier could not be confirmed, due to the erratic nature of the safety monitoring instrumentation. It has been commonly observed that instrumentations, while accurate under normal operating conditions, become unreliable under extreme physical conditions. This is due to being exposed to temperature, pressure or radiation levels that are beyond their safe operating range. It becomes stressful to verify with high confidence whether they are performing their desired functions, when the accident is already quickly escalating.

Some safety barriers may not be completely diagnosable, due to practical limitations. Fukushima's hydrogen leak from the containment vessel into the unit 1 reactor building or the hidden hydrogen leakage from unit 3 to unit 4 are examples of this. Certain operational deviations may remain hidden, despite considerable investment in monitoring features. This can be attributed to factors such as the type of barrier design, its location, nature of hazardous substance, system complexity, and monitoring instrument location.

Overall, there are several uncertainties associated with achieving the SDP's informational benefits. These arguments suggest that the SDP only partially satisfies the achievability criteria.

5.5. Relevancy discussion

Based on the findings from the achievability discussion, there are also reasons to question the relevancy. The SDP's relevance is determined by the value of information its objective provides. Acquiring the information on safety barriers' breach is clearly valuable on a standalone basis, But, when paired with DID, its relevance lies in improving the informed use of DID, which already requires barrier diagnosis. Then, we need to determine whether the SDP-motivated barrier diagnosis is more reliable, of higher quality or holds more real-time value to the operator managing a potential accidental event. If it improves the outcome more than when it is not implemented, its pairing with DID can also be justified economically.

The SDP's maximum informational value or relevance should be observed under accidental conditions, i.e. when the demand arises. Throughout the Fukushima accident sequence, the operators struggled to obtain information on safety barrier status to make accurate diagnoses. As already indicated, even if the SDP had been implemented, it would likely not have made a significant difference in uncovering the information, partly due to a limited scope. This undermines its ability to convey relevant information to improve DID's effectiveness.

The Fukushima accident is considered a man-made disaster, due to the failure of safety culture, management, regulators and government. If the SDP provided guidance on monitoring the weakening of these barriers, such a diagnosis would be material to improving DID's implementation and overall emergency preparedness. Then, it is possible that the accident's outcome could have been different and added business value. However, this is not the case, as the SDP does not address the diagnosability of such non-technical barriers (i.e. human and organisational barriers).

There can be outlier accidental scenarios, when safety diagnosability may not be relevant in bringing the hazardous plant state under control. For example, the Fukushima unit 3 operators could not have made use of the barriers' failure diagnosis, without the capability to act on this information. For a nuclear plant to be prepared for such situations, they need to regularly validate their design assumptions and invest in mitigatory/control measures. In addition, the questionable reliability of diagnoses received during emergency scenarios adds very little value, beyond placing attention to the quality of the information and whether one is compliant to the SDP. Under high-stress and hazardous situations, operators can lose the motivation to follow the SDP. As the SDP takes an extreme safety perspective without consideration for the actual economic benefits for the business, even management may lose enthusiasm for it.

Overall, SDP may be partially relevant, if it requires organisations to invest in its compliance without considering its true costs, benefits and associated uncertainties.

5.6. Timeliness discussion

As mentioned in 2.2, barriers of distinct levels and types are monitorable in different time frames. While the timely availability of diagnosis is undoubtedly critical, the SDP's overall objective is to maintain a superior barrier diagnosability, by making improvements period over period. This makes achieving the SDP's diagnosability an ongoing objective. Quantifying its time horizon is neither realistic nor logical. Therefore, this criterion is not applicable to the SDP and does not provide information about its usefulness.

6. Conclusions

The rationale behind the SDP is that a violation of its requirements increases the probability of an accident conditioned on an initiating event. SDP compliance means that, if situational awareness is degraded during system operation, it can be adjusted appropriately if, or when, the barriers are breached. This prevents the shrinking of the operator response window required to intervene

Table 4
SDP usefulness assessment result.

Rationality criterion	Criterion satisfied
Specificity	Yes
Measurability	Yes
Achievability	Partly
Relevance	Partly
Timely	Not applicable

effectively. This motivation is sound but builds on the premise that the information is obtainable.

A main argument for making the SDP information attractive is the insufficiency of the DID principle, but this is perhaps more a question of how DID is managed in this industry. With proper management, one could claim that the SDP would add limited value, as the relevant safety information, corresponding to what would have been provided by the SDP, is already available. It is an argument challenging the benefits of adopting two principles instead of just using DID.

A fundamental part of the DID principle is that, for the barriers to be reliable, management should recognise the importance of monitoring tools to diagnose the barrier and plant status. In particular, the defence layer at level 2 requires that operating experience is sent as feedback and that diagnostic tools record and announce information about faults in the control room. This is implemented by setting up instrumentation and control capabilities over the necessary ranges and through the use of digital technology of proven reliability [39]. This presents an element of redundancy, since diagnosability and feedback fall under DID.

Table 4 summarises the result of the SDP rationality assessment. From Section 5, it is concluded that SDP satisfies 'S' and 'M' and partially satisfies the 'A' and 'R', while 'T' is seen as inapplicable to this principle. The principle is clearly specific and measurable. Our discussion on its usefulness to the Fukushima nuclear accident case helped us derive general insights that strengthened the conclusions for 'A' and 'R'. These are important criteria that show that the SDP fails to completely satisfy these practical aspects. These are severe criticisms that can challenge the principle's usefulness, when employed to complement DID in the nuclear context, and it is a finding that can be generalised. This is because a specific and measurable safety principle has only limited usefulness if it is not completely achievable or lacks relevance to the business' safety. The Fukushima case study also shows that restoring the diagnosing capability, as per the SDP, would not have significantly improved the outcome.

On a standalone level, however, the situation might be different. It has not been our focus to assess this, and we recommend that future work should consider and conclude on the standalone benefits. We acknowledge that the SDP might show usefulness in combination with DID for some nuclear applications. Our conclusions, based only on this one accident, should not be generalised to cover all nuclear applications. Nevertheless, the SDP gaps pointed to are likely to apply to a wide range of applications, where the principle cannot be fulfilled, and might create a false sense of safety. Hence, we do not, on a general basis, recommend the implementation of the SDP for the nuclear industry.

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.

References

[1] J.H. Saleh, R.A. Haga, F.M. Favarò, E. Bakolas, Texas City refinery accident: Case study in breakdown of defense-in-depth and violation of the safety–diagnosability principle in design, Eng. Fail. Anal. 36 (2014) 121–133.
[2] J.H. Saleh, K.B. Marais, F.M. Favaró, System safety principles: A multidisciplinary engineering perspective, J. Loss Prev. Process Ind. 29 (2014) 283–294.
[3] ISO 12749-5. "Nuclear energy, nuclear technologies, and radiological protection – Vocabulary - Part 5: Nuclear reactors". 2018.
[4] R.L. Boring. "Adapting human reliability analysis from nuclear power to oil and gas applications.", No. INL/CON-15-35411. Idaho National Lab. (INL), Idaho Falls, ID (United States), 2015.
[5] H. Rosencrantz, K. Edvardsson, S.O. Hansson, Vision Zero–is it irrational? Transport. Res. Part A Policy Pract. 41 (6) (2007) 559–567.
[6] L.I.K. Sørskår, J.T. Selvik, E.B. Abrahamsen, On the use of the vision zero principle and the ALARP principle for production loss in the oil and gas industry, Reliab. Eng. Syst. Saf. 191 (2019), 106541.
[7] M. Fackler. "Report finds Japan underestimated tsunami danger". The New York Times. 1 June 2011. Available from: https://www.nytimes.com/2011/06/02/world/asia/02japan.html.
[8] International Atomic Energy Agency (IAEA). "Fundamental Safety Principles". Vienna: IAEA. 2006. Available from: file:///C:/Users/jts/Desktop/IAEA%20fundamental%20safety%20principles%20Pub1273_web.pdf.
[9] L. Chierici, G. Fiorini, S. La Rovere, P. Vestrucci, The evolution of defense in depth approach: A cross sectorial analysis, Open J. Safety Sci. Technol. 06 (2016) 35–54.
[10] United States Nuclear Regulatory Commission (US-NRC). "Historical Review and observations of Defence in Depth." Brookhaven National Laboratory, Upton NY (2016).
[11] ISO 1709. "Nuclear energy - Fissile materials - Principles of criticality safety in storing, handling and processing". 2018.
[12] International Atomic Energy Agency (IAEA). "IAEA Safety Glossary - Terminology Used in Nuclear Safety and Radiation Protection". Vienna: IAEA. 2018. Available from: file:///C:/Users/jts/Desktop/IAEA%20safety%20glossary%202018%20PUB1830_web.pdf.
[13] International Atomic Energy Agency, IAEA. Nuclear Security Series Glossary. Version 1.3. Vienna: IAEA. 2015. Available from: https://www.iaea.org/sites/default/files/18/08/nuclear-security-series-glossary-v1-3.pdf.
[14] International Nuclear Safety Advisory Group, INSAG. "Basic safety principles for Nuclear Power Plants". Safety Series NO. 75-INSAG-3, Rev. 1 INSAG-12, INSAG, Vienna, 1999.

[15] Western Europe Nuclear Regulatory Associations, WENRA, Reactor Harmonisation Working Group (RHWG). "Report- Safety of new NPP designs." WENRA, March 2013.
[16] United States Nuclear Regulatory Commission (USNRC). "NUREG-2150 A Proposed Risk Management Regulatory Framework". US Nuclear Regulatory Commission, Washington, DC. 2012.
[17] C.L. Smith. "Understanding concepts in the defence in depth strategy." In: Proceedings of the IEEE 37th Annual 2003 International Carnahan Conference on Security Technology 2003, pp. 8-16. IEEE, 2003.
[18] J. Ahn, C. Carson, M. Jensen, K. Juraku, S. Nagasaki, and S. Tanaka. "Reflections on the Fukushima Daiichi Nuclear Accident: Toward Social-Scientific Literacy and Engineering Resilience". Springer Nature, 2015.
[19] Tokyo Electric Power Company. "Fukushima nuclear accident analysis report." 2012.
[20] OECD/NEA. "Advanced Nuclear Reactor Safety Issues and Research Needs: Workshop Proceedings", Paris, France, 18-20 February 2002, Nuclear Safety, OECD Publishing, Paris.
[21] International Atomic Energy Agency (IAEA). "Safety of Nuclear Power Plants: Design". IAEA Safety Standards Series No. SSR-2/1, IAEA, Vienna. 2012.
[22] K.N. Fleming, F.A. Silady, A risk informed defense-in-depth framework for existing and advanced reactors, Reliab. Eng. Syst. Saf. 78 (3) (2002) 205–225.
[23] O. Akira. "Where was the weakness in application of defense-in-depth concept and why?" In Reflections on the Fukushima Daiichi Nuclear Accident, (2015): pp. 131-164. Springer.
[24] B. Li, M. Khlif-Bouassida, A. Toguyéni, Reduction rules for diagnosability analysis of complex systems modeled by labeled Petri nets, IEEE Trans. Autom. Sci. Eng. 17 (2) (2019) 1061–1069.
[25] J. Reason, Managing the risks of organizational accidents, Ashgate, Vermont, 1997.
[26] E. Bakolas, J.H. Saleh, Augmenting defense-in-depth with the concepts of observability and diagnosability from Control Theory and Discrete Event Systems, Reliab. Eng. Syst. Saf. 96 (1) (2011) 184–193.
[27] A. Paoli, S. Lafortune, Safe diagnosability for fault-tolerant supervision of discrete-event systems, Automatica 41 (8) (2005) 1335–1347.
[28] K. Edvardsson, S.O. Hansson, When is a goal rational? Soc. Choice Welfare 24 (2) (2005) 343–361.
[29] G.T. Doran, There's a SMART way to write management's goals and objectives, Management Rev. 70 (11) (1981) 35–36.
[30] J.T. Selvik, S. Bansal, E.B. Abrahamsen, On the use of criteria based on the SMART acronym to assess quality of performance indicators for safety management in the process industries, J. Loss Prev. Process Ind. 104392 (2021), https://doi.org/10.1016/j.jlp.2021.104392.
[31] Japan Meteorological Agency (JMA). "Information on the 2011 off the Pacific Coast of Tohoku Earthquake". 2015. Available from: http://www.jma.go.jp/jma/en/2011_Earthquake/Information_on_2011_Earthquake.html.
[32] International Atomic Energy Agency, IAEA. "The Fukushima Daiichi Accident, Technical Volume 1/5. Description and Context of the Accident". 2015.
[33] G. Brumfiel. "Fukushima reaches cold shutdown." Nature News (2011).
[34] K. Kurokawa, Fukushima nuclear accident independent investigation commission by the National Diet of Japan, Nippon Genshiryoku Gakkai-Shi 55 (3) (2013) 146–151.
[35] M.P. Cabasino, A. Giua, C. Seatzu, Diagnosability of discrete-event systems using labeled Petri nets, IEEE Trans. Autom. Sci. Eng. 11 (1) (2013) 144–153.
[36] G. Saji, Root cause study on hydrogen generation and explosion through radiation-induced electrolysis in the Fukushima Daiichi accident, Nucl. Eng. Des. 307 (2016) 64–76.
[37] U. Bielert, W. Breitung, A. Kotchourko, P. Royl, W. Scholtyssek, A. Veser, A. Beccantini, et al., Multi-dimensional simulation of hydrogen distribution and turbulent combustion in severe accidents, Nucl. Eng. Des. 209 (1–3) (2001) 165–172.
[38] International Atomic Energy Agency (IAEA), and the World Meteorological Organization. Flood Hazard for Nuclear Power Plants on Coastal and River Sites, IAEA Safety Standards Series No. NS-G-3.5, IAEA, Vienna (2003). (This publication is superseded by SSG-18 (2011)).
[39] International Nuclear Safety Advisory Group (INSAG). 1996. "Defence in Depth in Nuclear Safety". INSAG Series No. 10. Vienna: IAEA. 1996. Available from: https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1013e_web.pdf.