Safety Science 181 (2025) 106694

Contents lists available at ScienceDirect

Safety Science

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

Analyst behaviour and team processes during hazard analysis: The development of an observation protocol and initial results from evaluating HAZOP sessions

Per Øivind Braarud *, John Eidar Simensen

Institute for Energy Technology/OECD NEA Halden Human Technology Organisation Project, Norway

ARTICLE INFO

Keywords:
Hazard analysis
Risk assessment
Analyst behaviour and team processes
Expert-driven methods
Expert biases
HAZOP

ABSTRACT

Analyst behaviour and team processes are important factors in the quality of expert-driven hazard-analysis techniques and methods, such as HAZOP (Hazard and Operability Analysis) and STPA (System-Theoretic Process Analysis). Beyond the support provided by the analysis method, the literature suggests, for example, that facilitating creativity, awareness of human judgement limitations, and consistency of method application can substantially impact the completeness and outcome of the analysis. However, empirical research on these factors and their effect on hazard analysis is almost nonexistent. To address this gap, we (the authors) have developed an observation protocol consisting of 27 items on analyst behaviour and team processes. We developed the protocol based on a literature review and analyst interviews, utilising methods frequently employed to develop psychological tests. Two studies, with the participation of four analysts in each, found sufficient clarity and relevance of the protocol items and identified refinements and adjustments to the protocol. The analysts stated the protocol would be useful for analyst team self-evaluation in addition to its use in empirical studies. Future studies should verify our findings and could utilise the protocol in systematic studies of analyst behaviour and how it impacts the hazard analysis and the analysis outcome. Analyst teams could use the protocol for self-assessment in their professional development and as an indicator of the quality of their team processes.

1. Introduction

In safety demonstrations of digital instrumentation and control systems, hazard identification and hazard-analysis methods (e.g., Hazard and Operability Analysis [HAZOP], Failure Mode and Effect Analysis [FMEA], and System-Theoretic Process Analysis [STPA]) are essential for addressing the risk and reliability of digital systems (Ericson 2005; Smith and Simpson, 2011). How methods are applied (the process) is important when estimating the quality of the evidence (product), and for arguing in the safety demonstration.

Hazard-analysis methods, such as HAZOP and STPA, are usually performed by an ad hoc team consisting of multidisciplinary experts with extensive experience. Team leadership, team facilitation, and team process are key aspects in this complex setting (Swann & Preston, 1995; Rossing et al., 2010). Although methods for automated risk- and hazard analysis have been investigated (e.g., Suh et al., 1997; Garrett and Apostolakis, 2002), human-driven hazard identification will be essential for the analysis of the unique solutions involved in modernisation and new builds (Taylor, 2017). Research has investigated the choice of methods (Farooqi et al., 2022), compared methods, and investigated the performance of methods (Potts et al., 2014). Analyst behaviour and team processes have been discussed in the literature, and good practices have been recommended based on anecdotical observations (Baybutt, 2014; 2015a; 2018). However, empirical research on these factors and their effect on hazard analysis is almost nonexistent (Olsen et al., 2020).

In an effort to address the lack of empirical data, we (the authors) suggest an evaluation protocol for more uniform, systematic data collection. This effort will support the further development of hazard identification and analysis processes, including analyst team self-evaluation (Braarud & Simensen, 2020; Simensen & Braarud, 2020). The development of the protocol was based on methods commonly used to develop psychological testing and observation tools (DeVellis, 2017; Boateng et al., 2018). We identified protocol items from a literature review and from studies interviewing analysts. During the development, we tested the protocol with analyst teams performing a HAZOP on a nuclear safety fan system (Gran et al., 2020; Nelson et al., 2021).

* Corresponding author.
E-mail address: per.oivind.braarud@ife.no (P.Ø. Braarud).

https://doi.org/10.1016/j.ssci.2024.106694
Received 14 December 2023; Received in revised form 22 August 2024; Accepted 6 October 2024
Available online 11 October 2024

2. Development of the observation protocol

2.1. Approach

The first version of the observation protocol, aimed at initial expert testing and review, was developed using a structured approach consisting of the following steps: define purpose, perform domain analysis, generate items, develop a pilot version, expert reviews of the pilot protocol, conclude on the first protocol version (Murphy & Davidshofer, 1994; DeVellis, 2017; Boateng et al., 2018). Fig. 1 illustrates these steps.

2.2. Purpose and development of pilot version

We developed the observation protocol for use in analyst behaviour studies and for analyst team self-evaluation. The goal is that analysts should be able to apply the protocol independently and unassisted. This approach requires clarity and simplicity in the protocol, in addition to a straightforward description of its application.

2.3. Domain analysis, identification of dimensions, and item generation

We approached the domain analysis using an input, process, and outcome model (Ilgen et al., 2005). This approach resulted in an overall structure of a) input (team composition, facilities, and system documentation); b) analysis behaviour and team processes (facilitator behaviour, method application, evaluation, agenda, timing, workload); and c) outcome of the analysis (completeness and analyst learning).

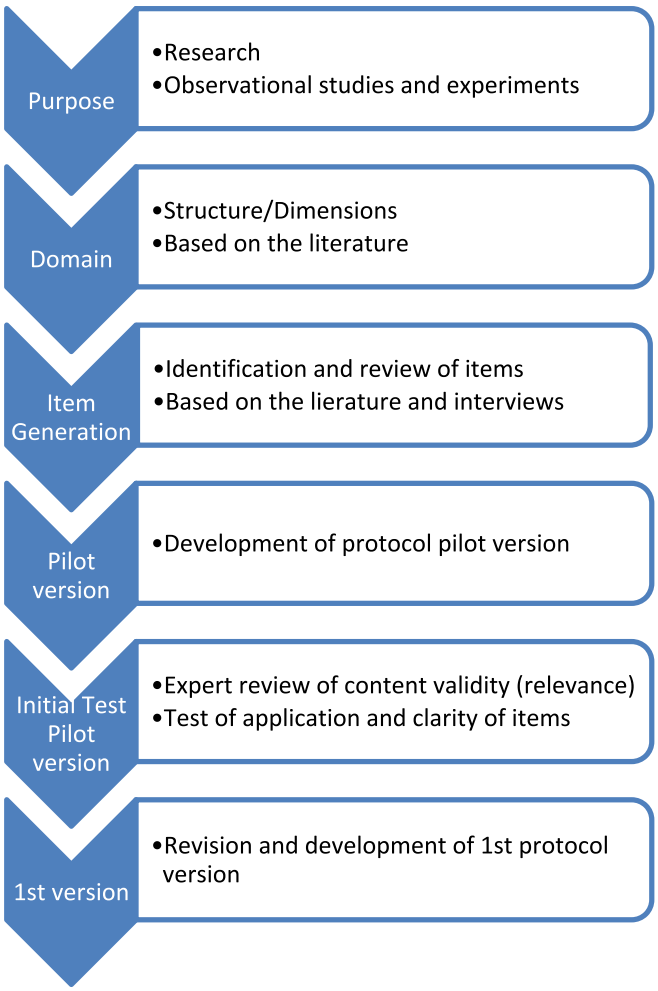

Fig. 1. Observation protocol development steps.

Based on a literature review and interviews with hazard-analysis experts (Karpati et al., 2014; Olsen et al., 2020), we individually reviewed the literature and the interviews and proposed items for each dimension. Similar items proposed by both authors were examined, and items for inclusion in the protocol was jointly formulated based on the literature. The following provides a brief summary of the basis of the items.

The HAZOP method can be applied in several ways (Baybutt, 2014; Baybutt, 2015a). A systematic and consistent application of the method supports completeness in hazard identification (Baybutt, 2014; IEC 2016). A HAZOP leader has the flexibility to, for example, change the order of the lead words according to the discussion or hazards, follow a hazard's path, choose timing constraints, and decide how to involve users. How the HAZOP is performed can impact, for example, the quality of results and analysis coverage (Swann & Preston, 1995; Baybutt, 2016; Baybutt, 2018). Creativity techniques, such as brainstorming, both individually and within the team, as well as shifting perspectives, support the identification of novel scenarios (Baybutt, 2014; Baybutt, 2015a). Furthermore, HAZOP terminology and the HAZOP process might be unfamiliar to many of the personnel without a background in hazard or safety analysis (Baybutt, 2015a; Rossing et al., 2015).

The evaluation of ideas is important in the team's analysis. Although creativity and team interaction might result in many potential scenarios and deviations, it is important to evaluate assumptions and uncertainties critically (Baybutt, 2014).

Clarifying the purpose of the analysis and setting the agenda help the analysis team follow the correct procedure without introducing secondary agendas (Whitty & Ford, 2009). Rossing et al. (2015) proposed a process to be followed by less experienced teams that support hazard focus.

HAZOP analysis can be intellectually demanding and time-consuming (Baybutt, 2014; Baybutt, 2015a): for example, the time set aside to focus on particular aspects, the time provided for the participants to think, and the time disposition according to what the facilitator or participants find the most important (Rossing et al., 2010; Munn, 2009; Baybutt, 2018). The complexity of the system under consideration can affect team members' focus, their ability to comprehend other team members, and their ability to identify hazards, which could lead to a high mental workload that reduces the quality of the team's analysis (Baybutt, 2018).

The facilitator needs to remain in control and manage experts, who may have strong opinions regarding the analysis. The HAZOP leader must control the analysis process while involving the team members to address the scope of the analysis and utilise the team's resources (Swann & Preston, 1995; Whitty & Ford, 2009). Schedule management involves focusing the team's attention on the important aspects of the system, and management of the team's workload helps maintain quality of work and avoid fatigue.

The outcome of a hazard analysis is frequently characterised by completeness regarding scenarios and the hazards identified (Rossing et al., 2010; Baybutt, 2014; Taylor, 2017). However, completeness might be difficult to judge for analysts without any specification of a complete solution to compare the team's results against (Taylor, 2017).

Munn (2009) suggests HAZOP analysis provides a good opportunity for participants to learn about the system under evaluation. Additionally, we found that learning about HAZOP method application could be an additional outcome of the HAZOP activity.

The documentation available for analysts can significantly impact the analysis results. As hazard identification and analysis are usually performed early in the system's development, information and documentation are often sparse and in progress (Taylor, 2017).

Facilities play an important role in supporting the participants and the process, such as joint surfaces, ease of access to information, and the welfare of the participants, as analysis can be lengthy (Swann & Preston, 1995; Baybutt, 2018; Olsen, 2020).

Team composition has substantial implications regarding the quality of the analysis. Of specific importance are competency in the analysis

method and the technical competences covered by the team members (Swann & Preston, 1995; Baybutt, 2014; Baybutt 2015b). Furthermore, a team that is too large can lose focus, dwell on an issue too long, or easily be disruptive. If all participants want to express their perspectives, this can lead to excessive discussion. However, a team that is too small might lack the necessary competency or perspectives to identify the right scenarios (Swann & Preston, 1995; Rossing et al., 2010; Trujillo et al., 2015).

For brevity, we present the dimensions and items of the protocol first version following the pilot test. Study 1 below describes the pilot testing, including the subsequent revisions, resulting in the first version of the observation protocol.

3. Study 1: Pilot test and development of the first version of the protocol

3.1. Pilot test method

The initial test and review of the protocol were performed by applying the protocol in a use-case study. The use case comprised a HAZOP (IEC, 2016) on a nuclear safety fan (Gran et al., 2020), which is a safety-relevant system (IEC, 2020).

3.1.1. Observation protocol

A pilot version of the protocol was applied in Study 1 to provide insights into the participants' understanding of and the relevance of the protocol items. For brevity, the final protocol is presented (see Section 3.9), and the implemented changes from the pilot version to the first version are provided in Sections 3.8.7, 3.8.8, and 3.8.9.

3.1.2. Participants' ratings of protocol items' clarity and content relevance

To become familiar with the protocol, each participant individually rated their experience of analyst behaviour and team processes following the HAZOP session. A seven-point Likert scale was applied (Likert, 1932). The participants were instructed to rate their experience such that, in their view, 1 represented an example of a very poor HAZOP on the given characteristic and 7 represented an example of a very good HAZOP on the given characteristic. Only the scale endpoints were labelled (Weijters et al., 2010). After completing the protocol rating, the observers were asked to rate the clarity and relevance (content validity) of each observation protocol item to assess the HAZOP process. Regarding content relevance, each item was rated individually on a four-point scale based on Davis (1992), consisting of 1: not relevant; 2: somewhat relevant; 3: quite relevant; and 4: highly relevant. This rating form enabled the participants to provide a comment on each individual item. Similarly, the participants were asked to rate the clarity of each item on a four-point scale ranging from 1 (unclear) to 4 (very clear). The rating form also allowed the participants to comment on each individual item.

3.1.3. Interview

After completing the individual ratings for item clarity and item relevance, a brief semi-structured group interview was conducted. The interview included the following leading questions: a) What were your experiences with answering the observation protocol? b) Do you have suggestions for additions, deletions, or adjustments to the protocol? c) What was your experience as a participant in this pilot study? d) What was your experience with the HAZOP analysis? For example, regarding the materials provided and the system being analysed. In addition to the questions, the participants were encouraged to comment freely on the observation protocol and the analysis performed. The interview considered each item of the protocol, asking for viewpoints on their clarity, completeness, need for revision, and relevance to the HAZOP sessions. The interview also collected thoughts on additional topics that should be included, such as protocol, as well as overall viewpoints about the protocol.

3.1.4. Study participants

Four participants fulfilled the roles of facilitator, secretary, and experts in the HAZOP sessions. One participant was assigned the combined roles of facilitator and secretary, while the other three had technical expert roles. The facilitator was a senior expert with experience in risk-and-safety analysis in several industry domains. The other three participants' backgrounds ranged from intermediate to senior expert on I&C (hardware and software), with experience in research and practical engineering. The study was reviewed and approved by the Halden Reactor Project Human Studies Review Committee. Informed consent was obtained from each participant. One study lead and an observer (the authors) monitored the HAZOP session. The authors also instructed and organised the protocol for data gathering and conducted the interviews in the study.

3.2. Hazard and Operability Study

A HAZOP study is a structured, systematic method for assessing complex systems to identify hazards (IEC, 2016). The HAZOP leader uses a set of guide words to support the study participants in providing their expertise on the system under consideration in a structured manner. The HAZOP method aims to add structure and systematism to subjective expert assessment. The system considered is evaluated by the group piece by piece, and the assessment is documented successively. As different aspects of the system are considered and different expertise is involved, the outcome of a HAZOP often depends on the strategy and approach of the leader and on group dynamics.

3.2.1. Safety fan system

The use case for the analysis sessions was a nuclear reactor safety fan system, the Halden Safety Fan (HSF; Gran et al., 2020). The HSF qualifies as a Nuclear Category B system – important to safety. The main function of the fan is to perform emergency air filtering and containment under accident conditions, such as radioactive leakage. The HSF system comprises two redundant filter stacks, including instrumentation and control, power supply, and human–machine interfaces. The system also contains both functional and non-functional requirements for its operation and performance, supporting both higher-level functional analysis activities and detailed technical analysis. As a use case, the HSF requires multidisciplinary competence for any detailed analysis and for integrated assessments (Swann & Preston, 1995). Although the use case is detailed, deliberate under-specifications of the technical components have been made to allow for variations in approach and solution.

3.2.2. Study 1 approach and HAZOP session

The HAZOP leader was briefed on the purpose of the study (protocol evaluation) and provided a briefing and documentation on the HSF. This process served two main purposes; first, the use-case material was provided to familiarise the HAZOP leader; second, the requirements and constraints for the HAZOP session were provided to the HAZOP leader, including the data protocols to be used for documentation. The study began with the study lead explaining the study purpose, the protocol evaluation procedure, and the access to and storage of participant evaluation data. The team was informed that 90 min were available for the analysis and that, following the analysis, an evaluation of the observation protocol would be performed. Following this briefing, informed consent was collected. In the HAZOP, the HAZOP leader was instructed to explain the HAZOP case – the safety fan – and that the HAZOP team should discuss a strategy for analysing the case. The session closed with a debriefing that asked about the participants' experiences of the study, including any experiences of discomfort, concerns about participation, or questions regarding the study.

3.2.3. Study conditions and equipment

All the participants were seated around a meeting room table. A large screen display (projector) provided a joint overview of both the case

documentation and the facilitator's notes during the analysis, and all the participants had a computer with the documentation. Additionally, a paper writing board, clean sheets, pens, and markers were available.

3.3. Results – Study 1

3.3.1. Item clarity

The overall rating for item clarity varied among the four participants. The median and the 25th and 75th percentiles (provided in the parentheses) for the four participants were 3.0 (2.0; 4.0), 4.0 (3.0; 4.0), 4.0 (3.0; 4.0), and 3.0 (3.0; 3.0). The median clarity across all the participants was 3.0 (3.0; 4.0). Of specific interest for the further development of the observation protocol were the items with a low clarity rating. Table 1 provides an example of an unclear item accompanied by participant comments to the rating, as well as an example of a clear item. Clear items frequently received no comments.

3.3.2. Item relevance

Overall, the participants viewed the protocol items as relevant for evaluating the human behaviour and team processes of HAZOP analysis. The median and the 25th and 75th percentiles (provided in the parentheses) for the four participants were 4.0 (3.0; 4.0), 3.0 (3.0; 4.0), 4.0 (2.0; 4.0), and 4.0 (3.0; 4.0). The mean relevance across all four participants was 4.0 (3.0; 4.0). Three items, all concerning learning from the HAZOP session, received a rating of 1 or 2 by at least two participants. Seven items received a rating of 1 or 2 from one participant only.

3.3.3. Participants' viewpoints

The detailed comments on specific items included proposals for reformulation to make items more understandable and relevant, and to remove less relevant items. The interview also resulted in proposals for additional items on the group climate and participation in the group's analysis regarding (1) to what extent the participant could participate in the analysis and to what extent the participant felt that her/his input was considered during the analysis; (2) the participant's experience of support and encouragement regarding participating in the group work; (3) the participant's experience of whether the group's effectiveness was effective; and (4) the extent to which discussions were useful for the group's analysis work. Table 2 summarises the topics the participants reported during the interview.

3.3.4. Pilot protocol revision process

A widely used approach for measuring content validity is the Item-level Content Validity Index (I-CVI; Martuza, 1977; Davis, 1992). The index was calculated for individual items (Polit & Beck, 2004). In our study, the I-CVI is the number of participants giving a rating of either 3 or 4, divided by the total number of participants. In other words, we dichotomised the four-point ordinal scale into relevant and irrelevant. A strong indication of item relevance was defined as I-CVI = 1.00, i.e., all four participants agreed on the item being relevant (Lynn, 1986). Items with low relevance were removed, e.g., I-CVI below 0.75 (Lynn, 1986; Davis, 1992). The interview results related to the low I-CVI items (relevance and clarity) were reviewed to understand whether low relevance was related to the item's clarity. Items of sufficient relevance but with low clarity were reformulated (Lynn, 1986; Davis, 1992), utilising the participants' comments. In addition, the items with high clarity were used as examples to formulate clear items.

Regarding the participants' proposed additions to the protocol, we first reviewed the existing items to check whether they should be updated rather than added. If this was not the case, we added a new item if the suggestion fit the overall dimensions identified in the literature. As such, we sought support in the literature to include new items based on the participants' comments.

Table 1
Examples of unclear and clear items and analysts' comments.

Item	Pilot item description	Median clarity	Participant comments
1	Application of method "as described and intended".	2.0	Had to make an interpretation of "*as described and as intended*". What is the comparison, what the facilitator intended, or compared with a baseline? Requires that people have experience with HAZOP or that an introduction is provided.
11	Facilitator performed the lead in managing the analysis process. Initiative / took the lead when needed, when expected.	4.0	No comments

Table 2
Topics from the participants' interview.

Topic	Content
Generally positive and relevant	The participants were positive about this type of self-evaluation and generally found the items relevant. They stated that an evaluation should always follow such a session.
Subjective aspects of evaluation	There is a subjective aspect – especially regarding the evaluation of the HAZOP process. One needed to interpret the items. The examples mentioned were as follows: What is the baseline for the evaluation? What is meant by "completeness" – is it "sufficiently complete"? For example, if people have an agenda, how am I supposed to know what they think?
Level of evaluation (group vs individual vs method)	What is to be evaluated at the group level vs individual level could be clearer, as well as what pertains to the group vs the hazard-analysis method.
Grouping of items	Items could be grouped more to make them clearer.
Addition of items	The participants wanted to add questions about whether one could express one's ideas, were able to participate actively, and whether the group members considered one's input to the analysis. The participants mentioned that if there are large groups, many people might be idle. They suggested adding more specific questions about whether one experienced support from the group during the analysis, as well as whether one was encouraged to participate in the group work.
The scale	It was difficult to apply the range of the scale.
Analyst experience	The interpretation of some questions depends on the analyst's level of experience with the hazard-analysis method.

3.3.5. Removal of items

Three items only, concerning learning from the session, received a relevance score of 0.5 or lower. The I-CVI were 0.25, 0.5, and 0.5 for the three items. In the interview, some participants stated that learning about the application of the HAZOP method and the system being analysed could be relevant in some settings. However, based on the low rating and the interview, we removed the learning items to focus the protocol on the most relevant content.

3.3.6. Modification of items

One comment pertaining to several items concerned the challenge of rating against what is viewed as an "external" criterion implied by the item phrasing: for example, questioning whether the HAZOP method was applied "as intended" or whether the identification of hazards was complete. These items were considered a comparison against an agreement or pre-definition of "as intended" and "complete". Establishing an agreement among participants regarding these terms when using the

protocol would have been ideal but impractical in terms of the time and resources required. The idea when developing the pilot version of the observation protocol was that the participants would evaluate the item against their own understanding of the criterion. Therefore, to make the items clearer and probably easier to answer, the items were modified specifically to evaluate the matter in terms of the participant's view (see, for example, the revised Item 1 presented in Table 3).

3.3.7. Addition of items

The participants suggested adding items regarding the extent to which the participant could participate in the analysis, the extent to which the participant perceived that her/his input was considered during the analysis, and the support and encouragement experienced for participating in the group work. The first two issues were covered in the pilot version from the perspective of evaluating the facilitator's contribution to the group process – the leadership of the group. One addition to the protocol concerned how each participant experienced their own involvement in analysis. We added a dimension labelled "Your involvement in the group work" with the following three items: To what extent were you able to participate in the analysis, to what extent did you feel that your input was considered by the group, and to what extent did you feel the group was supportive of your participation?

Furthermore, the participants suggested adding items regarding the extent to which one experienced that the group work was effective and whether the discussions were useful for the analysis. This item type captures the participants' overall subjective experience of the group's work. A dimension labelled "Group effectiveness" was added with the following two items: To what extent did you experience the group work to be effective in achieving the HAZOP goals, and to what extent did you experience that the group discussions were useful for the HAZOP analysis?

3.4. The first version of the of the observation protocol and procedure for use

The development of the pilot observation protocol and modifications based on the pilot testing resulted in the "first version" observation protocol. Table 3 presents the items in the first version.

A rating scale ranging from 1 to 7 was attached to each protocol item. An item was described by the item label and the description in Table 1. The scale endpoints were illustrated by "1: Example of very poor HAZOP on this characteristic, 7: Example of very good HAZOP on this characteristic". The procedure for the analysts' application of the protocol explained, "There are no 'correct' or 'wrong' answers to the items; the evaluation is looking for your viewpoints about the HAZOP analysis just completed." Furthermore, each protocol item included a field for analysts' comments. The procedure asked for comments on the item, such as elaborating on how the analyst perceived the item when rating it or asking for an explanation for the rating given.

4. Study 2: Test application of the protocol and investigation of analyst behaviour

Study 2 served two related purposes – testing user application of the observation protocol and investigating the utility of the protocol for the investigation of analyst behaviour. Furthermore, Study 2 included participant ratings of protocol item clarity and relevance. A similar approach as described for Study 1 was applied, with the following differences:

4.1. Observation protocol

Study 2 used the first version of the observation protocol (see Section 3.4 and Table 3).

Table 3
Observation protocol items and item descriptions.

	No	Item	Description
Method application	1	HAZOP application	We utilised the HAZOP method in line with my view of how the method should be applied.
	2	Flexibility	We were flexible in applying the HAZOP method (for example, use of lead words according to the discussion/hazard).
	3	Creativity techniques	We applied creative techniques when I thought they were needed (e.g., brainstorming).
	4	Terminology	We applied correct and consistent terminology.
Evaluation	5	Critical evaluation (Questioning attitude)	We critically evaluated and questioned assumptions of the analysis.
	6	Critical evaluation of probabilities and frequencies	We critically evaluated probabilities (scenarios) and critical evaluation of safeguards and considered uncertainty.
Agenda	7	Team's agenda	I experienced that one or more team member(s) advocated an agenda beyond the agenda determined by the HAZOP: for example, by raising topics beyond the HAZOP purpose.
Timing	8	Time available	To what degree did you think there was sufficient time available for the analysis (thinking, discussing, addressing important issues and challenges)?
Workload	9	Mental effort	How do you evaluate the effort you needed to invest in the HAZOP?
	10	System complexity	How do you consider the complexity of the system analysed (the components, systems, functionality – interconnections, number of parameters, size of system, etc.)?
Facilitator behaviour	11	Leadership initiative	I believe the facilitator performed the lead in managing the analysis process.
	12	Involving team members	I believe the facilitator supported (helped) the involvement of all team members.
	13	Schedule management	In my view, the facilitator kept an eye on the schedule and allocated time to the important parts of the analysis.
	14	Workload management	In my view, the breaks were adequate (e.g., periodically, or when quality of work deteriorated due to fatigue).
Effectiveness	15	Achieving HAZOP goals	To what extent did you experience the group work as effective in achieving the HAZOP goals?
	16	Group discussion utility	To what extent did you experience that the group discussions as useful for the HAZOP analysis?
Your involvement	17	Your participation	To what extent were you able to participate in the analysis?
	18	Group's response to you	To what extent did you experience that your input was considered by the group?
	19	Supportive group environment	To what extent did you experience the group as supportive for your participation?
Outcome	20	Hazards identified	In your view, to what extent did the group identify important hazards for the system?

(*continued on next page*)

Table 3 (*continued*)

	No	Item	Description
Documentation	21	Suitable detail of documentation	To what degree was information provided sufficiently detailed?
	22	Sufficiency of system documentation	To what extent was your need for information covered?
Facilities	23	Teamwork technology	To what extent did you find the work sufficiently supported by screens, posters, or other means to support group work?
	24	Work environment	How did you experience the level of comfort (e.g., level of noise, disturbances, room temperature) for the work?
Team composition	25	Technical competences	In your view, to what extent was the group's technical competence sufficient for the analysis?
	26	Method competence	To what extent did the group reveal sufficient knowledge and skills in applying the HAZOP method?
	27	Group size	In your view, how suitable was the group size for the HAZOP work (e.g., too many people resulting in inefficient communication, too few to establish good discussions)?

4.2. Study participants

Four HAZOP analysts participated (other than those participating in Study 1). The participants had different experiences of participating in HAZOPs, ranging from participation in one to several. They had backgrounds in engineering, informatics, security, and risk management. As in Study 1, Study 2 included two observers (the authors).

4.3. Study 2 approach and HAZOP session

The participants were provided with use-case documentation the day before the HAZOP session and were given up to three hours for individual preparation. The session was conducted as a hybrid meeting (MS Teams), with two participants online and two physically present. The participants were familiar with hybrid meetings and the MS Teams functionalities. Following the HAZOP, the participants were presented with a procedure on how to answer the protocols. The participants and observers responded individually to the protocol.

4.4. Study conditions and equipment

The equipment and support tools available were similar to those in Study 1. The online meeting solution provided the ability to share both single documents and entire screens with one user at a time. All the users had their cameras turned on. Live notes were available on a shared surface for all participants. Informed consent, according to the research institute's participant protection procedures, was collected from all participants before the session.

4.5. Participants' ratings of protocol items' clarity and content relevance

Study 2 employed the same procedure as Study 1 for participants' ratings of item clarity and item relevance (see Study 1, Section 3.1.2).

4.6. Session feedback

A structured group feedback session was organised, and feedback from each participant was gathered on the clarity and relevance of the protocol items.

4.7. Results of Study 2

Verification of the data provided by the HAZOP participants was performed by a team consisting of the two observers in conjunction with a third member who did not participate in the data gathering. The collected data were treated in two ways: the objective data (ranking score provided) were verified by one observer in two steps, and the subjective data were verified by observer agreement. Regarding the objective data (numbers), we ensured all fields were filled in before cross-checking the provided numbers with the subjective textual answers provided by the participants.

4.7.1. Review and verification of observation protocol data

To investigate the usability of the observation protocol, we reviewed the observation protocol data and utilised the participant interview for clarification. The review of participant ratings resulted in the clarification of four of the protocol items as follows:

Item 6 Critical evaluation of probabilities and frequencies. The analysts commented that they decided the evaluation of probabilities and frequencies was outside the scope of this 90-minute session. Therefore, this item was determined as inapplicable.

Item 7 Team's agenda. The review suggested that the application of the scale to this item was challenging. If interpreting the scale as low to high, a low rating could mean no agenda beyond HAZOP; if interpreting the scale as intended, no agenda would be an example of a good HAZOP focus: high rating (very good). Two analysts gave the item a rating of 1, but their comments made it clear that they meant "no agenda" beyond the purpose of the HAZOP. Therefore, regarding the analysis of analyst behaviour, the rating of 1 was replaced by 7 for the two analysts.

Item 9 Mental effort and Item 10 System complexity. All four raters gave the two items a low rating, meaning low effort and low complexity. The comments of two of the analysts suggest this interpretation.

Item 14 Workload management. Analyst 4 provided no rating for this item and commented, "The session was too short to build up fatigue." For analysis, the missing data were replaced by the mean of the three other analysts.

The review of the analysts' protocol comments revealed that some comments addressed HAZOP in general and some addressed the specific HAZOP session they had just conducted. Furthermore, some comments regarding a given protocol item included statements relevant to other items. To address the protocol items systematically, the analysts' statements were evaluated as follows: Three analysts jointly reviewed the analysts' comments and discussed 1) whether the statement addressed the item in question, 2) whether the statement addressed another item, or 3) whether the statement did not address a protocol item or other aspects of analyst behaviour or group processes of hazard analysis.

4.7.2. Item clarity an item relevance

The analysts generally perceived the items as clear and understandable. The median and the 25th and 75th percentiles (provided in the parentheses) for the four participants were 3.0 (3.0; 3.0), 4.0 (4.0; 4.0), 4.0 (3.0; 4.0), and 4.00 (4.0; 4.0). The median across the four analysts was 4.0 (3.0; 4.0). Although the lowest median for any item across all analysts was 3.0, the lowest clarity rating by any analyst was 2.0, and no item received a rating of 2.0 from more than one analyst.

The items were generally perceived as relevant for assessing human behaviour and team processes using hazard-analysis techniques and methods. The median and the 25th and 75th percentiles (provided in the parentheses) for the four participants were 3.0 (3.0; 3.0), 3.0 (3.0; 4.0), 3.0 (3.0; 4.0), and 3.00 (3.0; 4.0). The median across the four analysts was 3.0 (3.0; 3.0). Regarding low relevance, only one item received a relevance rating of 2.0 or less from at least two analysts. This one item, no. 2 on method flexibility, received a median score of 2.5. For the remaining items, the lowest median was 3.0.

Although the analysts overall perceived the items as clear and

relevant, they reported the observation protocol could be improved. The main statement included the following: a) the three items about "Your involvement in the group work" (nos. 17, 18, and 19 of Table 2) were a little repetitive and overlapping without adding precision to the ratings; b); b) better guidance was needed on how to answer the items regarding the different types of documentation (Items 21 and 22); and c) the rating and comments for some items depended a fair amount on the analyst's interpretation, e.g., Item 9 about mental effort.

4.7.3. Protocol utility: Rating of analyst behaviour

To investigate the utility of the protocol, we analysed the participants' assessments of the HAZOP session. Overall, the analysts characterised their behaviour and team processes positively. The individual mean ratings (the scale from 1 to 7 represented "very poor" to "very good") were 4.58 (SD = 1.14), 5.08 (SD = 1.29), 5.19 (SD = 1.20), and 6.04 (SD = 0.82), and the mean for all analysts was 5.22 (SD = 1.19). Correspondingly, the two observers were positive – average rating of 4.77 (SD = 1.15) and 4.95 (SD = 1.21) – with a mean of 4.86 (SD = 1.17).

The agreement among the analysts was evaluated using the intraclass correlation (ICC; Shrout and Fleiss, 1979). The ICC model applied was two-way random effects and absolute agreement. The agreement among the four analysts was ICC = 0.45 (F[25,37.3] = 2.09, p = 0.02), which can be classified as a fair agreement, according to Cicchetti (1994). The agreement between the two observers was ICC = 0.64 (F[21,21.6] = 2.76, p = 0.01) – a good agreement, according to Cicchetti (1994). The pairwise agreement between any analyst and any of the two observers ranged from ICC = 0.30 to ICC = 0.68, with a median ICC of 0.48. The relatively high agreement among the two observers compared with the agreement among analysts could be due to the observers' familiarity with the protocol item content and the application procedure. The modest agreement among the analysts suggests a substantial person-dependent element of the assessment.

Fig. 2 illustrates the mean rating of the four analysts and the mean rating of the two observers for each protocol item. The observers did not rate Items 17, 18, 19, and 24, because the items focused on the analysts' own behaviour. The vertical lines illustrate the minimum and maximum rating by any analyst or any observer.

The most positively rated characteristic was attending to the HAZOP purpose and not focusing on secondary agendas (Item 7, Team agenda). Other particularly positive characteristics were the leader involving team members and scheduling the work, the utility of group discussions, being able to participate in a supportive group environment, and the group size. The analysts rated their application of creative techniques positively, whereas the observers rated these aspects lower than the analysts. Less positive analyst ratings (but not poor according to the mean rating) were application of critical evaluation, time available for the analysis, and sufficiency of documentation. However, the observers rated critical evaluation higher than the analysts did.

4.7.4. Protocol utility – Qualitative insights

To illustrate an analysis of analyst behaviour, we conducted a thematical analysis of the analysts' and observers' comments and viewpoints, utilising their protocol as support in thematic analysis (Miles & Huberman, 1994; Braun & Clarke, 2006). The following key topics were identified:

Method application: The two facilitators (HAZOP leaders) chose different ways to apply the HAZOP method. One facilitator chose the components and the order in which to evaluate them, and the other facilitator listened to input from the participants about where and how to begin the assessment. The latter facilitator also chose to provide a table of evaluation parameters for the participants. The observers found this choice shifted focus away from evaluating the system/components and on to the HAZOP method itself. The observers noted this group seemed to struggle collectively with the chain of events and differentiating between the main functions and groupings of detailed functions. Therefore, we argue that managing expectations regarding method application and focusing on and agreeing on this application before the analysis session are important. This view corresponds with the observation that the participants reported it was useful to discuss the process

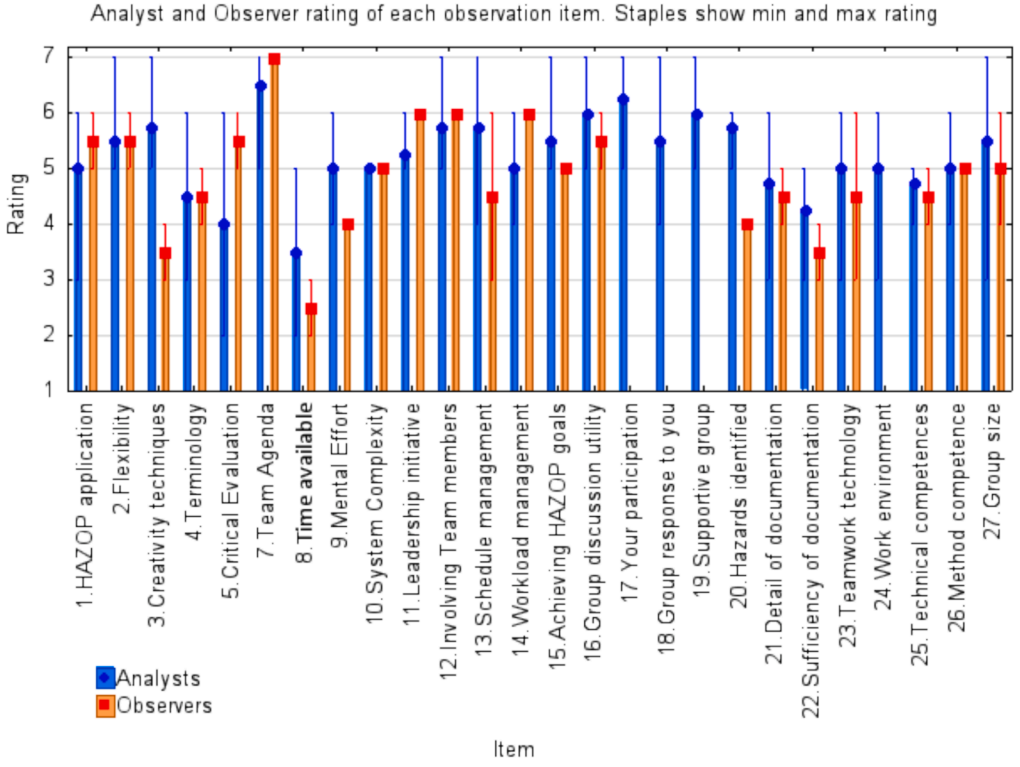

Fig. 2. Participants' and observers' ratings of analyst behaviour.

to be followed before the HAZOP session. Using the observation protocol to evaluate a HAZOP session can reveal ambiguities and different expectations among team members regarding the application of the method. Improvements on this topic can impact the team's ability to maintain focus on the HAZOP analysis, affecting the outcome of HAZOP sessions.

One of the participants highlighted a less rigid method application than expected, which sometimes made it difficult to track progress. Some HAZOP leaders follow a very rigid approach, covering all permutations of components and lead words, whereas others select some components and the lead words that make the most sense, then later extrapolate the results for other similar components and lead word combinations. The observers noted the same issue and mentioned that the analysis sometimes 'drifted away' from the target but noted that the 'drifting' resulted in hazards being identified, and the team usually returned to the original target. Both the analysts and the observers provided high scores for method application (Item 1, mean rating of 5.0 and 5.5, respectively) and flexibility (Item 2, mean rating of 5.5 for both), suggesting they perceived the method application as examples of a favourable approach.

The observers noted the team applied no "formal" creative techniques, but the team scored themselves high on this aspect (Item 3, mean rating of 5.75). We interpret this result as indicating the somewhat less rigid method application and the supporting team environment (see below) supporting creativity – which is a main goal of expertise-driven hazard analysis. However, the relatively low score from observers regarding creative techniques (Item 3, mean rating 3.5) is related to a formal creative technique not being applied.

The review of all the comments provided by the participants highlighted no issues regarding the consistent and correct use of terminology. The observers noticed some variation in the terminology used but saw no direct negative effect on the analysis process. However, the mean rating by both the participants and the observers was 4.5 (Item 4), which suggests room for improvement relative to the other topics investigated.

Facilitator behaviour: The observers reported the facilitator encouraged team member participation and that the leader asked for input from all the participants when they had been silent for a while or one of the analysts had been talking for a longer period. Both the analysts' and the observers' ratings suggest advantageous facilitator behaviour (Item 11, Leadership initiative, received scores of 5.25 and 6.0; Item 12, Involving team members, received mean scores of 5.75 and 6.0, respectively). However, to facilitate the discussions, the participants commented they would have preferred a clearer division of subsystems in the analysis process. This aspect highlights the potential for improving the technical elements of team leadership.

Group discussion: The participants rated the group discussions as highly positive (Item 16, Group discussion utility, mean rating of 6.0). The participants commented that the discussions were useful for coordinating the understanding of the system and to help provide useful inputs from the participants. The observers noted that all the team members' suggestions were positively met by the group. These positive group discussions were reflected in the analysts' ratings of the group's response to you (Item 18, mean score of 5.5) and the supportive group environment (Item 19, mean score of 6.0). However, the participants reported the discussions mostly occurred between the facilitator and the group and not so much within the group, and the observers noted that not all the participants were equally active in the discussion. The observers also noted this aspect might have been related to the session being performed as a hybrid meeting.

Participant involvement in the group work and supportive group environment: Overall, the participants rated positively the items regarding their involvement in the group work (Item 17, Your participation, mean scores of 6.25). One participant commented that the facilitator (HAZOP leader) explicitly supported the involvement of team members. The observers noted the facilitator motivated participation by asking for input from all members, including the quieter members. Furthermore, the observers noted the group members were, consequently, positive about each other's proposals and comments during the session.

Group effectiveness: The participants rated the achievement of HAZOP outcome (Item 20, mean rating of 5.75) as highly positive. However, the participants reported concerns about whether they sufficiently covered the scope and sufficiently addressed the most important system functionality. The participants also highlighted that the positive evaluation of the outcome was relative to the time available for the analysis. The observers rated the achievement of HAZOP outcome lower than the participants (Item 20, mean rating of 4.0 vs participants' mean rating of 5.75). A plausible explanation for the observers' lower rating is their knowledge about what hazards could have been identified and less attention paid to the limited time for the session.

Team composition: The mean ratings regarding group size were relatively high by both the participants and the observers (Item 27, mean rating of 5.5 and 5.0, respectively). Overall, the participants appreciated the teams' method competence (Item 25, mean rating of 4.75) and technical competence (Item 26, mean rating of 5.0). The observers shared this viewpoint and commented that a relatively small group seemed suitable for this type of online session.

5. Discussion

The reported work addresses the first steps towards establishing a systematic framework for investigating analyst behaviour and team processes regarding hazard and risk analysis methodologies. Study 1 developed a pilot version of an observation protocol for evaluating the human behaviour and team processes involved with hazard analysis. Adjustments were made following a pilot test. These changes concerned evaluating item clarity, item relevance, and usability, resulting in the observation protocol tested in Study 2. Overall, the analysts perceived the observation protocol as clear and understandable, and its content was relevant for assessing the human behaviour and team processes of hazard analysis. The participants in both Study 1 and Study 2 stated this type of evaluation would be useful following a HAZOP session. However, the participants mentioned some further improvements to the protocol. Study 2 suggested utility for use in research on analyst team behaviour, as well as utility for analyst team self-evaluation.

5.1. Clarity and relevance of the observation protocol

Study 1 found that the analysts viewed the pilot protocol version as relatively clear and understandable – the median clarity rating was 3.0 (the scale ranged from 1 to 4). The analysts' comments about protocol improvements suggested a viable approach for efficient self-assessment: to compare the analysts' experience of the HAZOP session with their expectations rather than comparing their experience with an external criterion. For example, it was feasible to evaluate whether the HAZOP method was applied according to the analysts' view of how the method should be applied. Evaluating whether the HAZOP method was applied as intended or in accordance with standards would require an explicit outline of the indented approach. This would be an interesting approach too, but it would be less efficient than the one chosen here. The approach chosen for this study, comparing experience with expectations, is similar to approaches utilised for self-assessment within human factors (Taylor, 1990; Hart, 2006; Braarud, 2021). Study 2 revealed that the analysts perceived the modified observation protocol as clear and understandable. The median clarity rating was 4.0. However, Study 2 identified a few further improvements to the protocol, but it also demonstrated the protocol is feasible and usable for rating analyst behaviour, which is an additional indication of the protocol being clear and sufficiently understandable for use.

Regarding relevance, in Study 1, the analysts viewed most of the pilot protocol items as relevant for investigating analyst behaviour: the median rating was 4.0 (the scale ranged from 1 to 4). However, the analysts provided a rating of less relevance for items concerning learning

from the session. In addition, the analysts proposed including items regarding analyst involvement in the team analysis. The Study 2 participants' median rating of relevance was 3.0, lower than the Study 1 median of 4.0. Although the protocol items were relevant, we believe that future research could investigate further improvements to the protocol. The Study 1 and 2 samples were relatively small, and larger samples might identify additional topics that should be included and provide more substantial evidence for the concrete specification of items.

5.2. Observation protocol utility

Study 2 found that the observation protocol was feasible for use in a thematic investigation of HAZOP analyst behaviour and team processes. Both the analysts' and the observers' ratings were used in the analysis, as well as the analysts' and observers' comments. The protocol provided a common framework for organising the observations and provided a domain-relevant structure for the thematic analysis. The example analysis presented in Study 2 suggests the protocol can be applied in empirical studies of analyst behaviour and for analyst team self-evaluation and learning processes. Since HAZOP mostly relies on the expertise and creativity of the analysis team, the quality of the analysis process is important for the HAZOP outcome. In this respect, the protocol ratings and the team self-evaluation can be used as elements in determining the quality of the HAZOP analysis.

Training guides and best practices for hazard analysis are plentiful, but they only provide the principles of the methodologies and describe what to achieve, not how, regarding, for example, process quality, confidence in results, and coverage. It is safe to assume that the application of this method varies substantially between companies. To identify and improve on best practices, a systematic, meaningful, and unified way to gather and evaluate both process- and product-aspects of hazard identification and hazard analysis is needed. We believe that the suggested protocol supports this data gathering and evaluation, and that industry and roles such as risk managers and analysts will be able to apply the protocol. Furthermore, the uniform, systematic way to conduct data gathering supports the assessment of method application across companies and industries.

As an example of protocol utility, the analysis in Section 4.7.4 found that a debriefing by using the observation protocol revealed different method applications that impacted the focus of the HAZOP session. Utilising the observation protocol as a guide in the session preparation serves as a reminder of the importance of establishing an agreement on the method application. This agreement supports the analysis team in focusing their expertise on the system under consideration and enables the facilitator to guide the team's HAZOP analysis by having the method application clarified before the session. Another example of utility is the participants' rating and evaluation of their degree of participation and contribution to the HAZOP analysis. For example, as reported in Section 4.7.4, the analysts provided high scores on these items (Item 17 rating 6.25, Item 18 rating 5.5, Item 19 rating 6.0). Protocol Items 17–19 on the ability to participate in the session and the team's responses to the analysts' contributions provide data relevant for evaluating the quality of the HAZOP. The extent to which different expertise has been able to and allowed to contribute to the output of the HAZOP can be a strong quality argument for the outcome of the analysis.

5.3. Limitations

Both studies investigating the protocol utilised small samples of four participants each. However, the second study generally confirmed the clarity and relevance suggested by the pilot study. As such, the second study provided replication evidence (Crandall & Sherman, 2016). Furthermore, both studies used participants with relevant education and experience from industry applications. The system under consideration was a realistic nuclear plant safety system (Gran et al., 2020), a use case featured in Nelson et al. (2021).

5.4. Future research

Study 2 collected the analysts' experiences and proposals regarding adjustments and improvements to include in the further development and refinement of the protocol. The procedure for applying the protocol for different purposes could be improved. The clarity and relevance of the protocol items could be investigated using larger samples. Furthermore, future research could investigate the need to add items to obtain a fuller understanding of the assessment of analyst behaviour and team processes; in other words, further investigation of the protocol's content validity is required.

Reliability regarding interrater agreement could also be further investigated (DeVellis, 2017; Boateng et al., 2018). The intraclass correlation suggests the agreement among analysts was moderate (Section 4.7.3). However, investigations regarding analyst team agreement and analyst–observer agreement could be conducted with larger samples. Furthermore, reliability could be approached by investigating the protocol's internal consistency.

The utility of the protocol could be further investigated using systematic studies of analyst behaviour. In addition, the protocol could be applied to controlled studies to understand better how analyst behaviour and team processes impact the HAZOP analysis and the analysis outcome. Furthermore, an improved understanding of human and organisational factors related to hazard analysis could also inspire future research on improving hazard and risk analysis methods and techniques.

6. Conclusion

The two studies developed and initially evaluated an observation protocol for analyst behaviour and team processes to understand better how human and organisational factors can impact the safety analysis methods used in the design of digital instrumentation & control systems. The studies found the protocol items were clear and understandable for analysts and relevant for application in studies of analyst behaviour and for analyst self-evaluation. The protocol can help document the quality of the hazard-analysis process and support analyst professional development. According to the authors' knowledge, this is the first protocol developed for collecting empirical data on analyst behaviour. The two studies employed small samples, but further studies utilising larger samples should verify our findings, and further development of the protocol would enhance its utility. We hope this work can inspire further research on analyst behaviour and team processes to improve hazard and risk analysis methods, procedures, and ultimately the outcome of such efforts.

CRediT authorship contribution statement

Per Øivind Braarud: Writing – original draft, Methodology, Formal analysis, Conceptualization. **John Eidar Simensen:** Writing – original draft, Project administration, Methodology, Conceptualization.

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.

Acknowledgements

This paper was based on research funded by the OECD Halden Reactor Project, Norway. Big thanks go to the participants in our two studies. Furthermore, we would like to thank our colleague Dr C. Esnoul at the Institute for Energy Technology, who was involved in the analysis of Study 2.

References

Baybutt, P., 2014. Requirements for improved process hazard analysis (PHA) methods. J. Loss Prev. Process Ind.

Baybutt, P., 2015a. A critique of the Hazard and Operability (HAZOP) study. J. Loss Prev. Process Ind. 33, 52–58.

Baybutt, P., 2015b. Competency requirements for process hazard analysis (PHA) teams. J. Loss Prev. Process Ind. 33, 151–158. https://doi.org/10.1016/j.jlp.2014.11.023.

Baybutt, P., 2018. On the completeness of scenario identification in process hazard analysis (PHA). J. Loss Prev. Process Ind.

Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quinonez, H.R., Young, S.L., 2018. Best Practices for developing and validating scales for health, social, and behavioral research: A primer. Front. Public Health 6, 149.

Braarud, P.Ø., 2021. Investigating the validity of subjective workload rating (NASA TLX) and subjective situation awareness rating (SART) for cognitively complex human–machine work. Int. J. Ind. Ergon.

Braarud, P. Ø., Simensen, J. E. (2020). HWR-1285 – "The development and initial testing of an observation protocol for investigating Analyst behavior and Team processes in Hazard Analysis", Report, OECD Halden Reactor Project, Norway.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qual. Res. Psychol. 3 (2), 77–101.

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. 6 (4), 284–290.

Crandall, C.S., Sherman, J.W., 2016. On the scientific superiority of conceptual replications for scientific progress. J. Exp. Soc. Psychol. 66, 93–99. https://doi.org/10.1016/j.jesp.2015.10.002.

Davis, L.L., 1992. Instrument review: Getting the most from a panel of experts. Appl. Nurs. Res. 5, 94–197.

DeVellis, R.F., 2017. Scale Development: Theory and Application, 4th edition. Sage Publications, Los Angeles, CA.

Ericson II., C.A., 2005. In: Hazard Analysis Techniques for System Safety. John Wiley & Sons. https://doi.org/10.1002/0471739421.

Farooqi, A., Ryan, B., Cobb, S., 2022. Using expert perspectives to explore factors affecting choice of methods in safety analysis. Saf. Sci. 146, 105571. https://doi.org/10.1016/j.ssci.2021.105571.

Garrett, C.J., Apostolakis, G.E., 2002. Automated hazard analysis of digital control systems. Reliab. Eng. Syst. Saf. 77 (1), 1–17. https://doi.org/10.1016/S0951-8320(02)00007-8.

Gran, B. A., Hauge, A. A., Simensen, J. E., Sarshar, S., Sechi, F., Gao, X., Sirola, M. (2020): HWR-1289 - "Halden Safety Fan – Context Description and System Specification", OECD Halden Reactor Project.

Hart, S., 2006. Nasa-task load index (Nasa-TLX); 20 years later. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 50. https://doi.org/10.1177/154193120605000909.

IEC (2016). International Electrotechnical Commission (IEC-61882 2016), "Hazard and operability studies (HAZOP studies) – Application guide", IEC 61882:2016.

IEC (2020). International Electrotechnical Commission (IEC-61226 2020), "Nuclear power plants - Instrumentation, control and electrical power systems important to safety - Categorization of functions and classification of systems", IEC 61226:2020.

Ilgen, D.R., Hollenbeck, J.R., Johnson, M., Jundt, D., 2005. Teams in organizations: from input-process-output models to IMOI models. Annu. Rev. Psychol. 56, 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250.

Karpati, P., Haige, A. H., Katta, V., Raspotnig, C. (2014): HWR-1112 - "Safety Demonstration and Justification of DI&C Systems in NPPs – Elicitation interviews with Regulators", Report, OECD Halden Reactor Project.

Likert, R., 1932. A technique for the measurement of attitudes. Arch. Psychol. 140, 1–55.

Lynn, M.R., 1986. Determination and quantification of content validity. Nurs. Res. 35, 382–385.

Martuza, V.R., 1977. Applying norm-referenced and criterion-referenced measurement in education. Boston, Allyn & Bacon.

Miles, M.B., Huberman, A.M., 1994. Qualitative data analysis: An expanded sourcebook, (2nd ed.). Sage Publications Inc.

Munn, A., 2009. Common Problems and Recent Trends with HAZOPs. IChemE Symp. 2009 (155), 130–133.

Murphy, K.R., Davidshofer, C.O., 1994. Psychological Testing: Principles and Applications, 3rd ed. Prentice-Hall, Englwood Cliffs, NJ.

Nelson, H.C.G., Kozine, I., Lundteigen, M.A., 2021. An integrated safety and security analysis for cyber-physical harm scenarios. Saf. Sci. 144 (0925), 105458. https://doi.org/10.1016/j.ssci.2021.105458.

Olsen S. A., Sarshar S., Simensen J. E., Reegaard K., Esnoul C. (2020). "Impact of Human and Organizational Factors Applying HAZOP: Results from a Systematic Literature Review and Interviews", Presented at Risk, Reliability and Societal Safety - ESREL 2020, Venice, Italy.

Polit, D.F., Beck, C.T., 2004. Nursing research: Principles and methods, (7th ed.). & Wilkins, Philadelphia, Lippincott, Williams.

Potts, H.W., Anderson, J.E., Colligan, L., Leach, P., Davis, S., Berman, J., 2014. Assessing the validity of prospective hazard analysis methods: a comparison of two techniques. BMC Health Serv. Res. 14, 41. https://doi.org/10.1186/1472-6963-14-41.

Rossing, N.L., Lind, M., Jensen, N., Jørgensen, S.B., 2010. (2010), "A functional HAZOP methodology". Comput. Chem. Eng. 34 (2), 244–253. https://doi.org/10.1016/j.compchemeng.2009.06.028.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86 (2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420.

Simensen, J.E., Braarud, P. Ø. (2020): HWR-1286 – "Human Behaviour and Team Processes of Safety Analysis Methodologies: Data Gathering and Results", Report, OECD Halden Reactor Project, Norway.

Smith, D. S., Simpson, K. G. L. (2011). "Safety Critical Systems Handbook - A straightforward Guide to Functional Safety, IEC 61508 and related standards", Elsevier Ltd., Third edition 2011, ISBN: 978-0-08-096781-3.

Suh, J.C., Lee, B., Kang, I.K., Yoon, E.S., 1997. An expert system for automated hazard analysis based on multimodel approach. Comput. Chem. Eng. 21 (Supplement 20), S917–S922. https://doi.org/10.1016/S0098-1354(97)87619-3.

Swann, C.D., Preston, M.L., 1995. Twenty-five years of HAZOPs. J. Loss Prevention, Process Industry 8 (6), 349–353.

Taylor, J.R., 2017. Automated HAZOP revisited. Process Safety Environ. Protection 111, 635–651. https://doi.org/10.1016/j.psep.2017.07.023.

Taylor, R.M. (1990). Situation awareness rating technique (SART): the development of a tool for aircrew systems design. In: A GARD-CP-478, Situation Awareness in Aerospace Operations. Advisory Group for Aerospace Research & Development, Neuilly Sur Seine, France, 3-1 to 3-17.

Trujillo, A., Kessler, W., Gaither, R., 2015. Common Mistakes When Conducting a HAZOP and How to Avoid Them. Chem. Eng. 122, 54–58.

Weijters, B., Cabooter, E., Schillewaert, N., 2010. The effect of rating scale format on response styles: The number of response categories and response category labels. Int. J. Res. Mark. 27 (3), 236–247.

Whitty, S., Ford, T., 2009. Is HAZOP worth all the effort it takes? IChemE Symposium Series No. 155, 143–148.