



Criteria for qualitative methods in human reliability analysis

Karin Laumann

Norwegian University of Science and Technology, Department of Psychology, 7491 Trondheim, Norway.

ABSTRACT

The qualitative analysis steps in Human Reliability Analysis (HRA) have received little attention in the HRA methods and in the general literature on HRA. The purpose of this paper is to investigate which quality criteria for qualitative research exist and if they can be useful as quality criteria in HRA. To find the criteria for qualitative research, a literature review was performed that covered a broad range of criteria for qualitative research. A thematic analysis was then used to sort and present the different criteria found in different papers and book chapters. Quality criteria for qualitative research were found in the following steps: presenting the background of the study, selecting and presenting a sample, qualitative data collection, qualitative data analysis, credibility/reliability checks, reflexivity and identification of possible bias, and evaluation of ethics considerations. It is discussed how these qualitative criteria could improve HRA and be guidelines for analysts and reviewers of HRAs.

1. Introduction

Swain [1, p. 301] states about human reliability analysis: ‘Human reliability analysis (HRA) can be used to estimate the quantitative or qualitative contribution of human performance to system reliability and safety.’ Kirwan [2, p. 7–10] suggests that an HRA consists of the following steps: ‘1. Problem definition, 2. Task analysis, 3. Human error identification, 4. Representation, 5. Human error quantification, 6. Impact assessment, 7. Error reduction analysis, and 8. Documentation and quality assurance.’ For the quantification part, human reliability analysis methods are developed. HRA methods often include general descriptions of (a) task types, which are tasks with certain characteristics defined by the HRA methods (in some methods called generic task types), and (b) performance-shaping factors (PSFs), with which the analyst must compare their specific data from tasks or scenarios. PSFs are also called performance influencing factors (PIFs) or error producing conditions (EPCs). A PSF is defined as: ‘an aspect of the human’s individual characteristics, environment, organization or task that’ specifically decrements or improves human performance, thus respectively increasing or decreasing the likelihood of human error’ [3, p. 177]. In HRA methods, the analyst should also evaluate how much influence the PSFs have on performance. For example, the Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) [4] and Petro-HRA, [5] use a restricted number of PSFs (compared to other methods); they also define specific levels of these PSFs to which they assign nominal multipliers.

Even if HRA often is described as a method to quantify the effect of human performance on task(s), most of the analysis behind the quantification is actually qualitative. For example, qualitative methods are used to define the problem (Kirwan’s [2] step 1), obtain information

about and describe the task and scenario (Kirwan’s [2] step 2), in human error identification (Kirwan’s [2] step 3). Qualitative methods are also the starting point in the quantification step (Kirwan’s [2] step 5), since the analyst usually collects information about tasks and PSFs and compares their data collected in an organization to task and PSF descriptions in the HRA method. All HRA methods I know include general descriptions of the elements included in the analysis. Therefore, the analyst has to use interpretation to compare the collected data from tasks or scenarios under analysis to the descriptions of PSFs, their levels and/or generic tasks described in the HRA method.

Qualitative data analysis has not received much attention within HRA compared to the quantification part. In the lessons learned from the US empirical study [6], one of the conclusions was that: ‘the methods do not provide sufficient guidelines or an explicit framework for analysts to conduct a structured and consistent qualitative analysis’ [6, p. 8]. If the qualitative part of an HRA is not performed well, the quantitative result might be unrealistic.

The background of this paper is that I have been working on the Petro-HRA project, which developed guidelines for HRA in the petroleum industry. During this project, we did some HRAs to test the guidelines. When discussing the results from the analysis, I thought that more guidelines than in the Petro-HRA, based on qualitative research methods, would have made the analysis more structured and transparent and facilitated its evaluation. Since my Ph.D. candidates and I have used qualitative research methods in several studies, I thought that advice for an HRA analyst based on criteria from qualitative research methods could be useful in an HRA. An HRA is usually not research; however, I would argue that criteria for qualitative research could be useful for improving the quality of an HRA analysis. In this paper, a literature review is performed to collect information about

E-mail address: Karin.laumann@ntnu.no.

<https://doi.org/10.1016/j.ress.2018.07.001>

Received 31 August 2017; Received in revised form 22 June 2018; Accepted 4 July 2018

0951-8320/© 2018 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

overall criteria for quality in qualitative research, and it is discussed how they could be implemented in a human reliability analysis. The research questions in this paper are: (a) Which quality criteria exist for qualitative research? and (b) How could these quality criteria be used to improve the quality of HRAs?

In HRA, there exist some guidelines about some of the steps, such as which qualitative data should be collected on PSFs [2,5] and how to do task analysis [7] as well as guidelines on how to define the purpose of the analysis [2,5]. However, the guidelines are descriptions of many approaches, and the analyst must select which one is most suitable for the analysis. Also, there are few criteria for what a good qualitative HRA is.

Some people seem to believe that HRA methods are scoring systems in which the analyst needs to apply little interpretation or reflective thought. In this view, one needs only to compare collected data to the description in the HRA method and decide a PSF, a PSF level and multiplier. However, others think that HRA methods are much closer to evaluation methods, where some guidelines are given and where the analyst does most of the interpretation. All descriptions of elements in HRA methods are general (and not context-specific); the analyst needs to interpret and reflectively compare the qualitative data collected in a specific setting to the descriptions of the elements in an HRA method. Since there seem to be different beliefs about what HRA methods are, this causes differences in the analyses that reflect these beliefs. My view is that HRA methods are closer to an evaluation method than a scoring system.

In this paper, overall criteria for qualitative research that would apply for all qualitative methods are described, so they could also be used as quality criteria for all the qualitative analysis steps included in an HRA. Further, there exist no guidelines the analyst can use in analyzing the qualitative data on the PSFs. Therefore, this paper will emphasize methods that could be used to give advice or guidelines to HRA analysts about how to perform the qualitative part of the analysis on PSFs.

To present quality criteria for qualitative research methods, some general information must be provided to give an understanding of the background and discussions about quality criteria in qualitative research methods. Hence, in the next section, the definitions of qualitative methods, paradigms in qualitative research and types of criteria within qualitative research are presented.

1.1. Definition of qualitative research methods

Malterud [8, p. 483] states: ‘Qualitative research methods involve the systematic collection, organization, and interpretation of textual material derived from talk or observation. It is used in the exploration of meanings of social phenomena as experienced by individuals themselves, in their natural context.’ It is difficult to describe what qualitative research methods are, because there are many different approaches with very different purposes. Often, all kinds of research that are not quantitative are called qualitative research. It is easier to describe what is typical for qualitative research than to give a specific definition.

- a) Qualitative research often involves analysis of textual data (such as an interview transcript, notes from observations and documents) [8,9,10,11].
- b) Data-collection often occurs in a naturalistic setting rather than in, for example, an experiment. [9,12].
- c) Qualitative research is usually concerned with understanding the research participant and getting his/her perspective. [13]
- d) Description and interpretation of data are necessary in qualitative research. [9,11,14].
- e) Qualitative research is concerned with the context in which the data occur [9,12,14,15,16, 17,18,19,20,21].
- f) Data collection and data analysis often occur iteratively [9,10,11].

1.2. Research paradigms in qualitative research

To understand discussions about qualitative research, it is important to know that there exist different research paradigms. These were described by Guba and Lincoln [22]. Other authors have presented different versions of these research paradigms. For simplicity in this paper, Guba and Lincoln's [22] much-used framework is presented. They define paradigm as ‘the basic belief system or worldview that guides the investigator, not only in choice of method but also on ontologically and epistemologically ways’ [22, p. 105]. They [22] define ontology as one's view of reality. Epistemology is about what kinds of knowledge we can obtain. Methodology is about how we can gain knowledge about the phenomenon we want to explore. Guba and Lincoln [22] described four paradigms that have different ontological, epistemological and methodological assumptions. The different paradigms are: Positivism, post-positivism, constructivism and critical theory. These paradigms are extensively described by Guba and Lincoln [22]; they are briefly presented next.

In the positivism paradigm, the main assumption is that one true reality exists and that this reality can be discovered by scientific methods. Further science builds objective, true knowledge or scientific laws. With a post-positivism paradigm, one assumes that reality exists, but that it is difficult or impossible to completely determine whether a true reality has been found. Within this perspective, ‘probable knowledge’ about an objective world can be obtained. In the constructivism paradigm ‘knowledge consists of those constructions about which there is relative consensus’ [22, p. 113]. In this perspective multiple realities can exist, and they can change. Moreover there is no objective truth that is separated from the knower. In critical theory paradigm, ‘knowledge consists of a series of structural/historical insights that will be transformed as time passes’ [22, p. 113]. The knowledge is co-constructed by individuals, a group and society and might be changing.

In a positivism paradigm, quantitative methods that include statistical tests of hypotheses and quantitative methods are applied. In post-positivism research, both qualitative and quantitative methods are used. With a constructivism and critical theory paradigm, qualitative methods are used to capture various interpretations of a subject.

I have not found that anyone has discussed the research paradigm that underlies HRA. HRA is concerned with probabilities for failures and successes on tasks. It fits with a post-positivism paradigm in which we cannot be completely sure that an objective true reality has been discovered; rather one only arrives at a probably true reality. A post-positivism view also fits with the use of qualitative data collection and analysis in HRA. Sometimes analysts ask for proof that the quantitative results from an HRA method are correct. Within a post-positivism paradigm, it is not possible to prove any results, however we can argue that the results are probably correct based on the information/data we have derived.

1.3. Criteria for qualitative research

There is a large literature that discusses what kinds of criteria should exist for qualitative research. Guba and Lincoln [23,24] developed criteria for qualitative research that are described as parallel to the traditional criteria of reliability and validity. These authors [23,24] claim that the overall criterion for qualitative research is trustworthiness. The parallel criteria they suggested were: credibility for internal validity, dependability for reliability, transferability for external validity and confirmability that is parallel to the criteria of objectivity. Guba and Lincoln's [23,24] criteria have had a tremendous influence on qualitative research, and they are still referenced much.

Morse [21] takes a different view, suggesting it is not necessary to apply parallel criteria for qualitative research. In her view, qualitative research should go back to the traditional research criteria for social science, such as rigour, reliability, validity and generalization. Meyrick [15] suggests transparency and systematicity as overall criteria for

qualitative research.

Whatever these overall criteria are called, they must be obtained with different methods in qualitative research than in quantitative research. This paper provides criteria for how to obtain these broad criteria.

Criteria for qualitative research have been extensively discussed. Rolfe claims that: ‘the literature on this issue can be broadly divided into three positions; those writers who wish qualitative research to be judged according to the same criteria as quantitative research, those that believe that a different set of criteria is needed; and those who question the appropriateness of any predetermined criteria for judging the quality of qualitative research’ [25, p. 304].

Even within the group of authors who think qualitative research should be judged by a different set of criteria than quantitative criteria, there is much discussion about which criteria should be applied. Patton describes seven types of criteria for qualitative research [26, p. 683–696]: ‘(1) traditional scientific research criteria, (2) social constructivism and constructivism criteria, (3) artistic and evocative criteria, (4) participatory and collaborative criteria, (5) critical change criteria, (6) system thinking and complexity criteria, (7) pragmatic, utilization-focused criteria.’

This paper searched for criteria for qualitative research that could be useful in an HRA. Often, authors do not necessarily indicate what types of criteria they present. However, in my judgment of the criteria selected for this paper, they are traditional science criteria, post-positivist criteria and social constructivism and constructivism criteria. Papers that describe other types of criteria were not found useful for the purpose of HRA.

2. Method

In this section, the methods used in this paper are presented. It has two parts: description of the literature search and how the data were analyzed.

2.1. Literature search

A literature search was performed in spring 2016, with search words such as criteria for qualitative research and guideline qualitative research and validity and qualitative research in the data bases Google Scholar and in PsycINFO. In the first selection of papers, the abstracts were read and, if the papers were judged to include criteria or guidelines for qualitative research, they were printed out and further investigated. The included papers dealt with criteria or guidelines for qualitative research overall methods and not for a particular qualitative method, such as grounded theory. Based on this search, 138 papers/book chapters were selected.

After the first selection of papers/book chapters, the papers/book chapters were investigated in more detail, and only those that included qualitative research criteria useful for HRA were included. After this search, 26 papers/book chapters were included. To ensure that I had included the newest criteria, a new search was performed for papers that were published in 2016 and 2017. This search resulted in three more papers/book chapters being included. To further look for newer papers, the references in the papers from 2016 to 2017 were investigated, and references that were newer than 2007 were included. From this search, 8 more papers/book chapters were included.

Altogether, 37 papers and book chapters were included in the data analysis. In the last papers included, there were no new criteria found, and since analyzing new papers did not yield any new criteria, the data were saturated – that I did not observe anything new or new variations in the searched criteria.

2.2. Data analysis

Data were analyzed, inspired by the thematic analysis steps of

Brown and Clark [27]. First, the papers were read and the different quality criteria described were marked, then the text that described the quality criteria was copied to a Microsoft Word document, and the criteria were given names. For the next papers, the criteria were compared with those already included, and if they were the same, the text describing those criteria were copied to the same criteria as the paper before. If the criteria were not included before, new criteria were developed in the Microsoft Word document. Sometimes, an author presented more than one criterion in a piece of information (sentence or paragraph), in such instances the information was copied to two or more criteria. After all papers and books were analyzed in this way, the content in the different criteria was read again, and a structure of the criteria was developed. In this process, some criteria that were similar were collapsed, and some criteria were split up into two criteria, because this gave better descriptions of those criteria. Finally, all papers were re-read and checked to ensure all criteria and references to the criteria had been included.

3. Results and discussion

In this section, the different criteria for qualitative research are presented, and their relevance for an HRA analysis is discussed. The papers in which the criteria for qualitative research were found are the following: [8–21, 28–50]. Table 1 summaries all the quality criteria for qualitative research methods found, in which analysis steps they are, and from which sources they come.

3.1. Background of analysis

Under this analysis step, different criteria are considered that should be attended to before the study/analysis starts.

3.1.1. Define and describe the purpose of the analysis [8,9,11–13,15,20,29,31,36–38,40,42,45,47,49]

In a qualitative method study, as for all kinds of research, it is important to describe a clear purpose and aim of the study [8,9,11–13,15,20,29,31,36–38,40,42,45,47,49], and clearly formulated research questions should be presented [9,12,47].

This should be reported in an HRA: An HRA analysis should describe its purpose and scope. Information and advice about how this should be done in an HRA can be found in Kirwan [2, p. 39–49].

3.1.2. Define and describe theoretical framework [8,9,10,14–16,18,19,28–30,32–37,40,42,45]

In a qualitative study, the researcher should clarify the theoretical background and former work that is relevant to the investigation [8,9,15,16,28,30,32]. One part of describing the theoretical framework is to identify the paradigm on which the research is based [9,15,18,29,36,37].

This should be reported in an HRA: In HRA different types of background information, such as which type of task analysis, error identification and HRA methods, etc., as have been chosen should be presented, and why that particular method was chosen should be discussed. In this part, there should be some discussion of the industry for which the method was developed and how the method fits, for example, in an industry other than for which it was developed. There might also be some other literature the analyst has used as background for the HRA. All this background information should be described. Describing the research paradigm on which HRA is based is not necessary for every analysis, since the paradigm in HRA fits with post-positivism.

3.2. Selection and description of sample

Different criteria that are about selecting and describing a sample are presented under this analysis step.

Table 1

Criteria for qualitative analysis, analysis step in which they occur and sources where the criteria were found.

Analysis step	Criteria for qualitative analysis	Source
Background of analysis	Define and describe the purpose of the analysis	[8,9,11–13,15,20,29,31,36–38,40,42,45,47,49]
	Define and describe theoretical framework	[8–10,14–16,18,19,28–30,32–37,40,42,45]
Selection and description of sample	Develop justification for the selection of the sample and decide how the sample is adequate for the purpose of the study/analysis	[8,9,11–13,15–21,28–30,35–38,42,45,47–49]
	Describe sample size and sample characteristics	[13,28,34,36–38]
Data collection	The analysts must consider which methods they should use to collect data	[8–16,18–20,28–30,32,34,36–38,40,42,44,45,47–49]
	Use and describe triangulation of data sources	[8,9,11,12,14,15,17–21,28–30,32–35,37–43,45–50]
	Describe context for data collection	[8,9,15,16,19,20,28,36,37,46–48]
	Describe duration of interview and observation and if multiple interviews and observations were performed	[28,34]
	Pilot or try out data collection method	[9,40]
	Describe method to record data and to process data	[8,11,14,28,37,46,47]
	Describe and justify change in focus and following up on surprises	[17]
	Obtain immersion with the data/have prolonged engagement in the field	[9,14,16–18,20,29,30,34,38, 39,41,43,46]
Data analysis	Describe methods to analyze data	[8,9,10,12,14,15,16,21,28,29, 34,36,37,40,42,44,45,47–49]
	Obtain and report a transparent process from data to conclusion/ findings should be grounded in the data	[8–16,19, 20,28–30,33–38,40,42,45,46,49]
	Include data from all subjects	[15,20,28–30,35,40,42,46]
	Weight the evidence	[17]
	Assess rival explanation(s)	[10,17]
	Perform and report a deviation case analysis/negative case analysis	[9,14,15,17,18,19,21,29,35,36, 38,41–43,46,47]
	Collect information about context and describe it in the report (thick descriptions)	[8,11,14–21, 29,30,33,36,38,40,41,43–46,48]
	Have coherence in the data presentation	[11,13,19,20,30,36,38,40,41,44, 49]
	Discuss impact and importance	[9,19,20,30,36,37,41,44,45,49]
	Discuss generalizing of research findings (transferability)	[13,14,16,17,19,20,36,37,42,43,45,47,48]
Credibility checks/reliability checks	Obtain and report feedback from participants/member checks	[8,9,13–21,28–30,34–42,45–48,50]
	Obtain and report feedback from peer debriefs and/or evaluation from external auditors	[9,12–14,17,18,21,29,38–40,43,46,48]
	Include multiple researchers and collaborators	[9,11,14,15,19,28,33,40,41,43, 45,46,47,48,50]
	Compare findings with other external factors	[13,40,45]
	Develop an audit trail	[9,11,12,14,15,17–19,29,33, 35,37,39,40,41,43,46,48]
Reflexivity and identifying possible bias (Criteria affecting the entire analysis)	Describe the researcher/analyst background and knowledge about the topic studied/analyzed and how this might affect results	[8,11–15,18–20,28, 30–32,36–49]
	Describe personal assumption that might affect the interpretation of data	[8,11–14,20,28–32,36–43,45–49]
	Describe relationship with participant and host organization	[9,11,28,30,36,37,41,49]
Ethics (Criteria affecting the entire analysis)	Evaluation of ethical considerations.	[9,20,30,36–38,41,44,47,48]

3.2.1. Develop justification for the selection of the sample and decide how the sample is adequate for the purpose of the study/analysis [8,9,11–13,15–21,28–30,35–38,42,45,47–49]

The researcher should decide what kinds of variations he/she wants to include from the data [11]. Usually it is not the number of participants that is important in qualitative research but more the adequacy of the sample in providing the necessary information on the research question [11,12,19,30]. The analyst should discuss the adequacy of the sample selected for answering the research question analyzed and give a rationale for why it was selected [15,16,28,30,36,37,49]. In qualitative data analysis, sampling is usually done stepwise. More participants are added after the beginning of data analysis if the analyst decides more information is necessary for answering the research question [8,11,49]. Thus the sample size is often not predetermined, as it is in quantitative research [8,21,48].

A focus group of subjects might also be chosen for data collection; if so the analyst should justify why a focus group was chosen and why those particular subjects were included.

The researcher could start with collecting information from participants that have most information about the topic and then, from this decide what kinds of variations are needed in the rest of the sample [11,12,29]. With qualitative methods, it is also possible to get information about people not included in the sample from those that are included [21]. In research the sampling of new informants stops at a point called saturation [12,48]. Saturation is achieved, and selection stops when new cases do not give new information, compared to prior participants [12,15,48]. To achieve saturation, information from

participants must be compared to each other, and as long as differences between participants give new insights on the topic, new participants are added [47,48]. When no new information is received from participants, selection of new informants stops [28,48].

In different guidelines for HRA, there is little information about who should be the basis of the analysis. In SPAR-H [4], the authors have claimed that it is an analysis of the average operator. This information does not help so much in deciding which operators or crew to include in an analysis, since the average operator probably does not exist or is impossible to identify at an early stage in the data collection. The analysts need to reflect on which and how many operators he/she needs to include for fulfilling the defined purpose of the analysis. With qualitative methods, it is also possible to include more operators during the data collection than initially planned, if the analyst finds this necessary.

This should be reported in an HRA: The analyst should give justification for the sample included in the analysis and discuss if the sample is adequate for the purpose of the specific HRA. Often the sample in an HRA consists of two or three operators. If the analyst includes two or three operators, he/she should discuss if the necessary information has been obtained to fulfil the defined purpose of the analysis. As stated above, with qualitative methods it is the amount of information and its quality that is important, not necessary the number of participants.

3.2.2. Describe sample size and sample characteristics [13,28,34,36–38]

The researcher should describe the sample and its basic characteristics [13,28,34,36–38]. This information gives a reader or reviewer information to see if enough variation was achieved to answer the

question. Information about how the sample was included in the study should also be given (e.g. was it the analyst who contacted the participants, or was it the organization, and was the participation completely voluntarily?).

This should be reported in an HRA: Also in an HRA, it is important to describe the sample that was included. The analyst should describe all characteristics of the crews included that might have an effect of the result obtained. Relevant characteristics that can be described in an HRA are age, gender, position in the organization, education, years of experience and amount of training. If variations among crews and operators that are important for the purpose of the analysis are not represented in the sample or data, these limitations should be described and discussed. An HRA should also describe how the sample was obtained.

3.3. Data collection

Different quality criteria that deal with data collection in qualitative research are presented under this analysis step.

3.3.1. The analyst must consider which methods to use to collect data [8–20,28–30,32,34,36–38,40,42,44,45,47–49]

The method chosen for data collection should be appropriate to answer the research questions [12,13,15,34,42,49]. The researcher should present the chosen method(s) and discuss whether the method (s) fit the research question [15,36,48,49]. The methods applied should be able to give the analyst the data needed to answer the research question.

The researcher should also describe in detail the method used to collect data [13,15,36,37,42,48,49] so that a reviewer can evaluate whether a good method was chosen [42,47]. With qualitative methods, the researcher might adjust the data collection method, depending on information they he/she obtain during data collection. Such changes should be described [15,37]. Sufficient detail about the methodology should be given so that it is possible to replicate the study [13,44].

The most used data collection methods in qualitative methods are interviews, document analysis, observation, photos and video recordings [12]. These methods should be planned; a data-collection protocol might be developed that describes the analyst's plan for the data-collection process and the methods used (interview questions, observational guide and documents that are included).

For interviews, the interview questions should be reasonable for the domain that is being investigated [14]. The questions should be open-ended, and the respondents should be able to elaborate on the answer [12,46]. The questions should not be misleading or try to elicit a certain response [46]. The interview questions should be presented in the report, and there should be a rationale for those questions [49]. Repeated interviews could be done to clarify meanings or to ask further about something the analyst discovered during data analysis. If this is done, it should be described in the final report. For interviews, focus group interviews could also be used [12].

For observation, an appropriate setting must be chosen [14]. The analysis should systematically collect data that are as concrete as possible on what was observed [14,40]. The data from an observation could also be captured with photos or video recordings. For document reviews, the analyst should describe the document(s) included in the analysis and establish the relevance for the study [14].

For HRA interviews, a semi-structured interview guide can be developed which consists of prepared questions, which are so open that interviewees can express and elaborate on their answers and where also follow up questions are allowed. This seems to be the most reasonable interview questions, since the analyst knows beforehand which topics should be covered in the interviews. However, the analyst should also be open to information not covered in the HRA method, for example, about a PSF that was not included. The analyst should also ask if there is more information about the performance on the task(s) that has not

been covered in the interview to check if some information was missed. Some examples of interview questions for the PSFs in Petro-HRA could be found in the Petro-HRA Guideline [5].

This should be reported in an HRA: The analyst should describe the data-collection methods in detail. A justification for the specific approach used should be provided. For transparency, the analyst should include the interview guide in the report.

For observational data collection, the HRA analyst needs to describe the exact setting for observation(s) and how data were derived from them. It could, for example, be by photos, videos or notes. The analyst should also describe in detail which documents were used and how and why they were included.

3.3.2. Use and describe triangulation of data sources [8,9,11,12,14,15,17–21,28–30,32–35,37–50]

Collecting data applying several methods and from different perspectives are recommended within qualitative research methods and is called triangulation [8,9,11,12,14,17,18,21,30,35,39,45,47,50]. Triangulation can have two purposes. It can be used as a reliability check to explore if more than one method yield the same data and the same conclusions [20,21,40,41,42]. Triangulation can also be applied as a method to investigate data derived by different methods or from different sources, it enables consideration of those data from different angles or perspectives [8,20,21,39,42,47,50]. This is the approach most qualitative researchers recommend [19,50]. This view on triangulation is also called crystallization [20,50]. It often gives the researcher a wider and more comprehensive understanding of the topic being analyzed [11,17,19,20,21,42,47,50].

Triangulation of data sources is also important in an HRA. The analyst could collect data on performance-shaping factors with multiple methods and approaches. For example, most HRAs include a PSF about available time or time pressure. To understand how time affects the performance of the task or scenario, the analyst could; (a) ask the operators to give an overall description of how much time they would use on a task and how much time they think they have available; (b) develop a timeline with the operators and discuss how long each task will take; (c) simulate how the task is done in the control room and measure the time used; (d) include documents that describe how much time the operators have available to complete the task. All these ways to explore the PSF with respect to time will yield different types of information about time that could be useful in evaluating how its different aspects might affect performance. Also, to gaining information on the PSFs from different perspectives (operators, trainers, management) might give different sources of information that also could be used to evaluate the PSFs' level (in, for example, SPAR-H [4] and Petro-HRA [5]).

This should be reported in an HRA: The analyst should reflect on which data collection methods are most suitable to obtain specific information in the HRA analysis for example on a specific PSF. If the same type of information is obtained with different methods this should be described, since this give support for the findings. It should also be described if different data-collection methods yielded different results. How this could be dealt with is further discussed under weighing the evidence described in session 3.4.4.

3.3.3. Describe context for data collection [8,9,15,16,19,20,28,36,37,46–48]

The analyst should select and describe the context for the data analysis [8,9,15,16,19,28,36,37,46,48] and also discuss how the context could have an effect on the data collected [9,19,28,46]. A question that could be discussed is how much of what is observed is affected by the analyst's presence [46]. Another question could be if there are some reasons for the subjects not telling the analyst everything about an issue or being biased. The analyst should also present a rationale for the particular setting that was chosen for data collection [16]. As well the analyst should also indicate if non-participants were present during the data collection and how they might affect the data

[28].

In some HRAs, the analyst has performed an HRA within an office and has not actually seen the environment where the task or scenario is performed; in some cases, in the design phase, this is the only option. Some of the PSFs might be challenging to evaluate without being in the actual setting where the task(s) are performed; this should be discussed.

This should be reported in an HRA: The analyst should describe the setting(s) in which the data were collected and indicate who were present during the data collection. The analyst should discuss possible limitations in obtaining good data caused by the setting, people present or other organizational issue.

3.3.4. Describe duration of interview and observation and if multiple interviews and observations were performed [28,34]

The analyst should describe the duration of interviews and observation, since this might provide information about the amount of data obtained [28]. For the same reasons, the analyst should also describe whether repeated interviews or observations were performed [28,34].

This should be reported in an HRA: The analysts should describe the duration of data-collection (interviews and observations) and if repeated interviews were performed.

3.3.5. Pilot or try out data-collection method [9,40]

The researcher could try out or pilot data collection and data recording methods. This could provide the researcher some training in performing the data collection and an opportunity to test out questions and structure for the data collection [9,40]. This process might be particularly useful if the analyst has little experience with interviews and observations.

This should be reported in an HRA: HRA data are challenging to obtain; collecting data for methods such as task analysis seems to be particularly challenging. The analyst should have some practical training in collection data for HRA before the real data collection starts. If he/she has little experience in performing HRA it should be described in the report how data collection methods were piloted or tried out. For experienced analysts this might not be necessary.

3.3.6. Describe method to record data and to process data [8,11,14,28,37,46,47]

The method used to record data should be described [28,37]. With qualitative data analyses, the interviews are usually audio-recorded and transcribed, since this more accurately captures what the participants actually say, than the mere analyst notes [11,14,28,46]. If audio recording are not included the researcher should give reasons why not [28]. The analyst should describe how data from observation were obtained, such as with videos, photos or notes during or directly after the observation.

This should be reported in an HRA: An analyst should describe in detail how data were obtained. Interviews should be recorded and transcribed. Observations should preferably be recorded with photos, videos and notes directly after the observation. Reasons for not using data recording methods should be presented.

3.3.7. Describe and justify change in focus and following up on surprises [17]

A researcher should have a plan for data collection. However, with qualitative data, the researcher might obtain information that to some degree, changes the focus during data collection. If such changes in data collection were made, reasons for this should be described [17].

This should be reported in an HRA: Also in an HRA, change from the original data-collection plan and reasons for doing this should be described. For example, if a PSF not included in the HRA methods seems to have a large effect on performance, one might consider methods to obtain more information about this PSF, beyond what was anticipated in the original data-collection plan. This should be described in the

report.

3.3.8. Obtain immersion with the data/have prolonged engagement in the field [9,14,16–18,20,29,30,34,38,39,41,43,46]

A criterion that is much discussed in qualitative research methods is prolonged engagement, which means that the analyst has collected a sufficient amount of evidence or has been in, for example, an organization for a sufficient amount of time to understand the topic investigated [14,17,18,39,41]. Also, being in contact with the organization for a longer time might increase the participants' trust in the analyst [17,41]. Prolonged engagement might help the analyst identify what is relevant and what is not relevant information [18].

This should be reported in an HRA: In an HRA, the analyst should discuss if the method used and the time spent in the organization yielded a sufficient amount of data and understanding to fully analyze the task, scenario and performance-shaping factors. At least for some HRAs, one might question if the analyst has collected a sufficient amount of information to analyze the tasks and all the potential performance-shaping factors. If an analyst, for example, just interviews one operator and has not done any field observations, one might question whether there is enough evidence to analyze for example the PSFs. However, the amount of data collected could depend on the purpose of the analysis, so the analyst should present some discussion about the amount of data collected and the time spent on collecting data and being in the organization. From this, a judgment can be made if it was sufficient for the purpose. If the analyst presents his/her view on this, a reader or reviewer could make an independent judgement about the sufficiency of the data collected.

3.4. Data analysis

Under this analysis step, different quality criteria on data analysis are presented.

3.4.1. Describe methods to analyze data [8–10,12,14–16,21,28,29,34,36,37,40,42,44,45,47–49]

The data analysis method selected must also be appropriate to answer the research question [8–10,12,14,15,16,21,28,34,36,37,40,42,44,48,49]. During data analysis, the qualitative information obtained is usually condensed and structured in ways that could serve as the basis for the analyst's interpretation [9,12,14,28,34,37]. This process should be systematic and planned, with a procedure that the analyst follows [8,16,48]. The analyst should present details about how the data were analyzed and discuss if the method was appropriate [8,15,16,37,42,48,49]. The presentation of data analysis should be so thorough that a reader or reviewer could follow all the steps that were taken [9,49]. In addition, the presentation of the findings should be consistent with the data analysis approach chosen [49]. Usually it is not enough to say that a technique or method (for example, thematic analysis) was used. The analyst should tell exactly how the data were analyzed [8,15,16,42,48,49]. The description should also present who performed the data analysis [37].

In HRA, there are general guidelines for how to do some part of the qualitative data analysis, such as task analysis [7] or error identification [2]. When the analyst uses these techniques in an HRA, he/she should describe exactly how the task analysis or error identification was performed.

I have not found any description of a method the analyst could use to analyze the qualitative data on PSFs. One exception might be that there are some descriptions about how to do, how to develop and how to analyze a timeline or how much time the analyst would use on a task [5]. Most HRA analyses I have seen do not include such a description of how the qualitative data were analyzed. A suggestion for a technique that can be used to structure the data on PSFs from the textual data (such as interview transcripts, documents and notes from observation) is template analysis, described by Brooks and King [51,52,53]. Brooks

and King [51, p.1] define template analysis as: ‘a technique for thematically organizing and analyzing textual data.’ Template analysis is a flexible technique that also allows the use of a *priori codes* [53] developed before the analyst starts to analyze the data. King [53 p. 431] states: ‘Coding is the process of attaching a label (code) to a section of text to index it as related to a theme.’ Most qualitative data analysis includes some form for coding of the data [9,10,21,36]. Coding is used to reduce and/or sort data [32]. In template analysis, the analyst investigates the data and marks each meaningful text related to the topic that is described as a top-down investigation of the data. In an HRA, the PSFs could be the *a priori codes* in a template analysis. However, an HRA analyst should also be open to including PSFs that affect performance that are not defined before the analysis starts. For example, if weather is an important PSF for a specific task and the analyst uses a method where weather is not listed as a PSF, he/she could add this to the list of codes (or themes). The data could also be sorted in a hierarchy, where different aspects of PSFs would be coded as a second level. The analyst could give the same segment of text more than two codes, for example, if the participant talks about more than one PSF in one sentence or paragraph. The analyst could also code connections between or links between PSFs and develop new codes for the links. After template analysis or coding, the HRA analyst must investigate and interpret the information coded for each topic or theme and compare the information to the PSFs and PSF levels in the HRA method.

Sometimes HRA analysts do not understand that it is necessary to interpret the information collected on the PSFs and just try to map directly something one operator said to the PSFs definitions and levels, for example, in SPAR-H [4]. Then there is no way that a reader necessarily knows why the analyst thought this information is evidence for that particular PSF or PSF level. The data collected must be interpreted, and the analyst must describe how he/she interprets the data to fit with data in the HRA method, for example with a PSF and a PSF level in SPAR-H [4]. The HRA methods are developed to support the analyst's interpretation of the data collected. However, HRA methods are not specific enough to all contexts and all aspects of a PSF, so interpretation must be performed by the analyst in order to compare the data from the specific context to the more general description given in the HRA methods. A reviewer needs to see a description of these interpretations to understand the analyst's interpretations.

This should be reported in an HRA: The analyst should present in detail how he/she structured and analyzed the data. It is not enough to say that a specific method was applied. The analyst should describe exactly how the data analysis method was used. The analyst should also describe his/hers interpretations of how the collected data fit with the information in the HRA method. For example, in SPAR-H [4] the analyst should describe how they determine that the collected information fits with a specific PSF or PSF level.

3.4.2. Obtain a transparent process from data to conclusion/findings should be grounded in the data [8–16,19,20,28–30,33,34,35–38,40,42,45,46,49]

An important criterion in qualitative research is that the reader should be able to follow the process from data to conclusion [8,13,15,19,28,30]. The analyst should present raw data, as well, the interpretation made from the raw data [11,13,17,19,36,37]. In qualitative data analysis, this is called grounding the interpretation of the data [11,16]. To do this, the analyst should audio-record interviews and transcribe them. Tong, Sainsbury and Craig [28, p.356] states that ‘Generally, audio recordings and transcriptions more accurately reflect the participants’ views than contemporaneous researcher notes.’ Recording and transcribing the data is an important part of all qualitative research methods, and it is a basic assumption underlying all qualitative data analysis. From the interviews, the analyst should include quotes from the informants on which the interpretation was based [8,9,14,19,28,29,32,34,36,37]. With observation, a concrete description of what was observed and how it was interpreted should be presented [40]. Observations could also be documented with photos or

videos. Also, for documents, the exact words should be included [37] along with the analyst's interpretation. The purpose of having a transparent process from data to conclusion is that someone else could review and discuss the correctness of the interpretation made from the data [9,11–15,28,29,30,36]. If more people can agree on the interpretations of the data, intersubjective/inter-assessor agreement could be said to occur, which could support the validity of the findings [31]. If two people disagree about the interpretation, one of the interpretations is not necessarily incorrect; it could mean that it is possible to interpret the data in more than one way. However, if the raw data are presented, this would also make it possible to discuss and argue about which interpretations are most appropriate. If too little of the original data (quotes from interviews, concrete descriptions from videos or quotes from documents) are included, another person cannot independently evaluate the interpretations made by the analyst [9].

Sometimes in an HRA, only the raw data are presented without any interpretation [6] as evidence for a PSF and a PSF level. The analyses where the analyst tries directly to match what the operators said to a level or to PSFs and where there is no sign that any thinking or interpretation was done usually makes the analysis very poor. The US empirical study concluded [6] ‘However, the analysts differed in the extent to which the interview information was used in their analysis. Some analysts tended to rely on direct input from the interviews, while others tended to rely on their own analysis and judgement with interview information as a supplement.’ The US empirical study [6] also seems to assume that analyses based on the analyst's interpretation is necessary. However, sometimes the analysts also only present their interpretations and not any raw data. Then it is not possible for a reader or reviewer to evaluate the interpretations or see from where the interpretation came. The analyst should present both raw data and the interpretation of the raw data.

This should be reported in an HRA: The report should include both raw data (that is structured) and the analyst's interpretation of how the raw data fits with the information in the HRA method (such as task types, PSFs or PSF levels). Also, with task analysis, it is important not just to present the output of the task analysis (for example, a diagrammatic hierarchical task format); this makes it impossible for a reader or reviewer to evaluate the analysis, since all the information on which the analysis was based is not there. The final report should also include raw data from the transcriptions of the interviews or concrete descriptions from observations that were used in developing the task analysis diagrams.

3.4.3. Include data from all subjects [15,20,28–30,35,40,42,46]

All cases should be included in the data analysis and in the report [15,29,42,46]. The analyst should not just present findings that support his/her conclusions [29,46]. Pilnick and Swift [42] and Lewis [46], claim that the analyst has a tendency to find more easily evidence that confirms his/her belief. An active search for data that do not confirm the analyst's expectation could increase trust in the analysis. Also, if the information from different types of subjects deviates this information should be presented.

This should be reported in an HRA: In the HRA report, the analysis should include information about how the operators or crews are different. HRA has been said to be an analysis of the average operator. However, it could be equally important for the interpretation of the PSFs and how data are analyzed to include an analysis of differences. The data from different informants should be compared to each other and the differences should be described. It should also be described how the analyst dealt with differences within the crews when the quantitative HRA analysis was performed. If there is different information from different groups of subjects, such as trainers, operators and management this should be described in the report. It should also be described if and how this was taken into account in the quantitative part of the analysis. In addition, in task analysis, different operators and trainers might present different ways that the tasks could be

accomplished. All of these different approaches should be included in the report. The report should not present just the most typical way the task is done, since this might obscure that differences exist. However, the analyst should describe if and how this was taken into account in the quantitative part of the analysis.

3.4.4. *Weight the evidence* [17]

The analyst must compare different types of data to each other and determine the best evidence [17]. In this process, the analyst must decide which are the best data to be used [17].

This should be reported in an HRA: With qualitative data, the analyst might obtain inconsistent evidence from different sources or different subjects. In this case, the analyst must weigh the evidence and discuss which seems to be the best or the most reliable. For example, if the operators claim that a task can be performed in a very short time, but an analysis includes an evaluation of how much time each task step takes, this approach might give a longer time estimate. In this case, depending on the question analyzed, the analyst must decide which evidence is the most reliable.

3.4.5. *Assess rival explanations* [10,17]

The analyst should investigate likely rival explanation of their data [10,17] Onwuegbuzie and Leech [17] suggest that it is difficult for analysts to detach themselves from their initial interpretation of the data. If it is not possible from the data to decide which one of two rival explanations is more correct, more data should be collected [17]. An example of this in HRA is if the analyst with a cognitive task analysis has found that the operators have the knowledge to perform a task, but they claim that they do not have any experience or training on the task. In this situation, it is difficult to select a level for experience and training. The analyst though could go back to the operators and ask where he/she got the knowledge, since this will help him/her select a level for the PSF experience and training.

This should be reported in an HRA: The analyst should evaluate and discuss rival explanations/interpretations of the data. For example, in an HRA it is sometimes difficult to decide on which PSF or PSF level some information from interviews should be placed. In this case, the rival explanation should be presented and the selection should be discussed.

3.4.6. *Perform and report deviation case analysis/negative case analysis* [9,14,15,17–19,21,29,35,36,38,41–43,46,47]

This criterion is related to the criterion; ‘include data from all subjects’ described in section 3.4.3. In deviation case analysis the analyst looks for cases that contrast with interpretations he/she made [9,14,15,17–19,21,29,35,36,42,46,47]. The negative case analysis could help in producing new and alternative interpretations [42,46]. One example of deviation case analysis is to use cases that, in some meaning, are extreme [17,46].

Deviation case analysis is about checking that all relevant information is taken into account, especially the information that deviates from the analyst's interpretation. Also in HRAs, there might be cases or information that deviate from the analyst's interpretation of the data. The analyst should search for such information to confirm whether the interpretations are correct, if they should be changes, or if some more analysis is required. Deviation case analysis is related to the purpose of the analysis and the analysis should include all information related to the purpose. The deviation case analysis assures that the analyst has searched for alternative interpretations of the tasks or the PSFs.

This should be reported in an HRA: The analyst should account for how a deviation case analysis was done, what was found, and if something was changed, based on the analysis.

3.4.7. *Collect information about context and describe it in the report (Thick descriptions)* [8,11,14–21,29,30,33,36,38,40,41,43,44–46,48]

To present information about context in which the data occur is an

often-mentioned criterion for qualitative research. There should be sufficient description of the contexts or settings in which the data occur, so the data are meaningful to the reader [11,17,19,20,40,46]. In qualitative data analysis, this is called thick descriptions [16–18,20,21,41,45,46,48]. The researcher must present the entire picture in which the data exist [20]. In research this is also important, because this description is used to evaluate to which contexts/settings the result can be generalized. [15–18,40,48].

Generalization of data to other settings is not relevant for HRA. However, an understanding of the context or setting in which the task or scenario occurs is important. If an analyst only presents pieces of information about each PSF, it is difficult for a reader to understand the entire picture of the data and how and why the PSFs affect performance.

This should be reported in an HRA: The analyst should present contextual information that is important for the reader in understanding the qualitative and quantitative part of the analysis. The analyst should describe the context in which the tasks or scenarios occur. Also, information about the organization, the crews (how many crews, how they are organized, etc.) and the setting in which the task(s) or scenario(s) occur would increase a reader's or a reviewer's understanding of the scenario, the tasks, and the PSFs.

3.4.8. *Have coherence in the data presentation* [11,13,19,20,30,36,38,40,41,44,49]

Related to the former criterion is the criterion coherence in the data presentation. This criterion state that information should be presented as coherent and integrated, which is the opposite of presenting unrelated pieces of information [11,13,19,20,30,36,38,40,41,44,49]. The data should be presented with meaningful coherence [20], and it should be explained how different pieces of data relate to each other [11].

Sometimes in HRA analysis, only a small amount of information is presented on each PSF, without describing the context in which the data occur or how the PSFs interact. In the real world, PSFs do not exist in isolation; oftentimes more than one PSF is present. The analyst must interpret if the influence is already represented to some extent by another PSF. In such instances there is a risk of double counting the same effect and so only one of these PSFs should be included. However, in the qualitative analysis, the analyst must consider how the PSFs occur together and in the context (crew, scenario, task, organization) in which they occur.

This should be reported in an HRA: the analyst should present a coherent picture of the data. For example, he/she needs to present a coherent picture of the PSFs and describe how they are interrelated, (rather than pieces of information on each PSF, where they seem unrelated). The purpose of this information is to give the reader/reviewer the whole picture of the tasks and the PSFs.

3.4.9. *Discuss impact and importance* [9,19,20,30,36,37,41,44,45,49]

In qualitative research, one should describe and discuss the implication of the research findings for further research, methods, knowledge, and/or practice [9,19,20,30,36,37,41,44,49].

This should be reported in an HRA: Also, in an HRA, the implication of the qualitative and quantitative results for the task and scenario should be presented. The analysis should not only present a quantitative likelihood for failure or success on a task or scenario as a result of the analysis, but also discuss; a) the qualitative insights on general tasks and PSFs that are the basis for the quantitative likelihood obtained and b) how these might be changed to improve performance (error reduction).

3.4.10. *Discuss generalizing of research findings (transferability)* [13,14,16,17,19,20,36,37,42,43,45,47,48]

For generalizing of research findings, the term transferability is often used in qualitative research. This criterion is about if and how research findings from one context can be used in another context [36].

For qualitative research, evaluating to what degree and how the study findings can be transferred to other situations or contexts and presentation of the study contexts [thick description], as described in Section 3.4.7, are important. This gives a reader insights as to what other contexts the study could be generalized to or not [13,14,16,36,41,43,48].

Generalizing findings to other contexts is usually not important in an HRA, since an HRA investigates one particular context, and the findings are usually not transferred to other contexts. However, in the development of HRA methods, transferability of findings and conclusions is important. Transferability should be clearly described in an HRA method, (which data and expert judgements generic task probability and PSF multipliers are based on), so the analyst can make a judgement about how much this information is applicable for his/her analysis.

This should be reported in an HRA: Generalization is not relevant to report for every HRA unless the analysis is included in a database or another development of HRA.

3.5. Credibility checks/reliability checks

Several quality criteria for qualitative research include methods to check the correctness of data and conclusions. In the next sections, several of these methods are presented.

3.5.1. Feedback from participants/member checks [8,9,13–21,28–30,34,35–42,45,46–48,50]

Feedback from participants could consist of (a) obtaining feedback from the participant on the correctness of raw data, for example, the transcripts [14,15,17,18,21,28,29,40,41,46,47], (b) getting feedback from the participants on the interpretation of the findings and conclusions, [9,14,15,17–21,28,29,34,36,38–42,46,47,50] or (c) going back to the informant with questions during the data analysis to confirm some interpretations [14,17,20,29,38,50]. It is discussed in the literature on qualitative method what exactly is possible to get the informants feedback about. An analysis is often a synthesis of data from many participants, and one participant could not provide information on other participants' data, only his/her own [21,50]. For the feedback on the interpretations and conclusions, usually the analyst has more knowledge about the topic studied and can therefore interpret the data differently from the participants [19,21,30]. However, with feedback from informants, one can determine if some misinterpretation or misunderstanding has occurred or if the participant would like to add some information [9,19,46]. The researchers can also use the feedback to evaluate their own interpretations [42]. The informant's feedback could be seen as another source of data that could be further interpreted, and thus increase the comprehensiveness of the study [20,42,47]. When the analyst uses feedback from participants or member checks, he/she should report what kind of information the respondent received, the feedback that was received from the participants and what he/she did with it, for example, where something changed in the analysis [16].

Obtaining feedback from participants on the qualitative data in an HRA is very important. HRA is a complex analysis, and lack of data and misunderstandings can happen. The analyst should be an expert on HRA and probably have more knowledge than the operators on how PSFs affect performance. However, the operators have most knowledge about the specific systems, tasks and work environment under analysis. Feedback from participants on data and interpretation of data should always be obtained, since this might reduce the risk for misinterpretation.

This should be reported in an HRA: The HRA report should contain descriptions of what kind of information the participants got, the feedback the analyst received and what he/she did with the feedback (if something was changed based on the feedback).

3.5.2. Obtain and report feedback from peer debriefs and/or external auditors [9,12–14,17,18,21,29,38–40,43,46,48]

Peer debriefs are about having another researcher review the data and final report [9,13,14,17,18,21,29,38–40,46,48]. The peer debrief might provide another perspective on data collection, data analysis, interpretations and conclusions [9,17,18,29,38,46,48]. A peer debrief might question and challenge the analyst's view and help him/her explore his/her interpretations [17,28]. Peer debriefs could also be an external evaluation that might serve same purposes as a peer debrief [14,17,21,38,45]. Peer debrief could occur during or after the data analysis, or both. Peer debrief should be a systematic evaluation of the work [18]. The results of the peer debrief and how the results of this activity were used should be included in the report [21]. Morse [21] warns against uncritically changing the interpretation based on peer debriefs, since it is the analyst who has collected and analyzed the data, and is thus closest to the data. When peer debriefs are used, the analyst should listen to and evaluate alternative points of view, but should take full responsibility for the final result [21].

This should be reported in an HRA: Peer debriefs could also be done in an HRA. A peer analyst might question data and interpretation done by the HRA analyst. This should be done in a systematic manner. The report should include a description of how the peer debrief/external audition was done, and if and how it changed some results. Also, in an HRA, it is the main analyst who is responsible for the report. The peer debrief is for questioning, for example, if enough data have been collected and the interpretations of PSFs and PSF levels; however, it is the HRA analyst who makes the final decisions on the interpretation presented in the report.

3.5.3. Include multiple researchers and collaborators [9,11,14,15,19,28,33,40,41,43,45–48,50]

Multiple researchers serve many of the same purposes as peer debriefs, but here, more than one analyst is involved in performing all steps of an analysis [14,40,45]. With more than one researcher, the data can be coded and interpreted by different analysts to check if two analysts will code or interpret the data in the same way [9,11,15,19,28,46]. If the data are not interpreted in the same way, the analyst should describe how such inconsistency was resolved [48].

HRA analysis could be performed with a team of analysts. To have multiple coders might be too time-consuming for each HRA. However, it might be useful for research on HRA methods to explore if analysts code the data in the same way. Also, it might be interesting to explore how consistently the data are interpreted. With a more systematic process, first coding and then interpretation based on the HRA method, it might be possible to explore where possible differences between analysts occur.

This should be reported in an HRA: The report should describe if and how more than one analyst was involved in the HRA. If more than, one analyst codes the data the process should be described. It should also be described how disagreements (if any) were solved.

3.5.4. Compare findings with other external factors [13,40,45]

Qualitative data could also be compared to other external data to check for consistency [13].

This should be reported in an HRA analysis: If it is possible, the results from an HRA should be compared to other data. The results could be compared to, for example, accident reports or to other HRAs performed on other tasks and scenarios, in this manner determining if the finding makes sense compared to other analyses. The comparison with other data should be described in the report.

3.5.5. Develop an audit trail [9,11,12,14,15,17–19,29,33,35,37,39,40,41,43,46,48]

One criterion for qualitative research very much related to all the above criteria is the creation of an audit trail [9,11,12,14,15,17–19,29,33,35,39,40,46,48]. This consist of writing

down all information about decisions and actions that were taken during the preparation of analysis, data collection and data analysis [9,12,18,19,46]. This process may make the decisions more structured and planned [9]. It also may make it possible to keep a record of information during the analysis [9]. With an audit trail, it is possible for a reviewer to evaluate the process that occurred during the study [48]. An audit trail should be made available for a reviewer [9,46].

An audit trail could also be useful in HRA since, in HRA analysis, there is much information of which to keep track, and it would make the analysis more planned and structured. It also gives much information to a reviewer about the process that was followed and why different decisions were made.

This should be reported in an HRA: The other criteria describe what kinds of information the analyst should present from the qualitative data analysis. The audit trail is used to keep track of all the information. The audit trail should not be presented in the report in itself; however, it should be possible for a reviewer to assess it, if requested.

3.6. Reflexivity and identifying possible bias (criteria affecting the entire analysis)

The criteria sorted under this main topic are different biases; both internal factors to the analyst and external factors might affect interpretations of the data [8,28,30,41,48,49]. Being open about such factors is called reflexivity, and these are important criteria for qualitative research [8,11,12,14,15,18,19,30,31,36,38,39,41,42,44,45]. Identifying such factors is assumed to reduce biases [11,12,21,36,42,45,46,49].

3.6.1. Describe the analyst's background and knowledge about the topic studied/analyzed and how this might affect results [8,11–15,18,19,20,28,30–32,36–49]

The researcher should describe his/her own background, which is important for the research questions, since this might have an effect on the analysis performed [8,11,19,28,32,37,41,42,48]. This could be, for example, personal characteristics, interest, age, sex, education and experience with the topic and methods [28].

This should be reported in an HRA: The reports should contain information about the analyst's educational and professional background. The analyst should describe knowledge about, training on, and experience with, HRA. The analyst should also indicate his/her experience with the industry in which the HRA is performed. Factors such as these might have a large effect on how methods are used and how data are interpreted.

3.6.2. Describe personal assumption that might affect the interpretation of data [8,11–14,20,28–32,36–43,45–49]

One criterion for qualitative research is that the researcher reflects on and is open with his/her personal assumptions, values or beliefs on the research topic, since this might influence his/her interpretation of data [8,11–13,20,28–32,36–43,45–49]. Ideally, the researcher should be aware of and try to limit these influences in the analysis [8,11,12,20,32,36,41,45]. If the researcher is open and explicit about these and reports them, it is also possible for a reader or reviewer to assess whether they might have affected the interpretations of data [8,11,28,32,41,46,48]. If the researcher and reviewer are aware of the researcher's/analyst's assumptions and beliefs, it might make it easier to consider alternative interpretations [28,30].

This criterion is also important in an HRA. An analyst should reflect about his/her thoughts and beliefs regarding a task or how likely operators are to fail on the task before he/she does the analysis; it might be possible for an analyst to thereby see if they affected the interpretations. Sometimes when reading an HRA report, it seems like it is the analyst's personal belief about an appropriate error rate that is the basis for the interpretation performed on the PSF, rather than interpretation of the data itself. If the analyst has been open about his/her

personal beliefs, it might make this more transparent both for the analyst and the reviewer.

This should be reported in an HRA: The analyst should present his/her assumptions before he/she did the analysis, the basis for the assumptions and discuss if these expectations affected the results.

3.6.3. Describe relationship with participant and host organization [9,11,28,30,36,37,41,49]

The nature of the relationship between the participant and the organization should be described, since this relationship might have an influence on how the data are interpreted [9,28,36]. If this relationship is described in the report, a reader or reviewer could make his/her own judgment about how these relationships might effect the data [9].

An example of how this relationship might have an effect in an HRA is that based on earlier discussion with participants or with the organization, an analyst might have developed a strong expectation that there will be a low likelihood of failure on a task; this belief could affect how the data are interpreted. Such expectations should be described.

These types of influences might be everything from very explicit to subtle, and the expectation might also be only an interpretation from the analyst that a certain result is expected. However, if these types of influences are presented, it might be easier for a reviewer and others to understand the analyst's interpretation of the qualitative data. It is not possible to remove these kinds of influences; however, the analyst should be aware of them and discuss how they might have influence data analysis, so a reviewer can judge if they had an influence the [11]. Ideally, the analyst should maintain independence from political and relational influences on the analysis; however, this is not possible to completely achieve [9]. If the information or relationship with a participant and host organization could affect the data and results, this should be described and discussed.

This should be reported in an HRA: The analyst should describe his/her relationship with the participants and organization and discuss how this might have affected how the data were interpreted.

3.7. Evaluation of ethical consideration (criteria affecting the entire analysis) [9,20,30,36–38,41,44,47,48]

Ethical consideration is a criterion that was mentioned by many of the papers on qualitative research criteria [9,20,30,36–38,41,48]. There exist ethical standards for research, such as the 'do not harm' principle [9,20,38], respect the participant [20,38], confidentiality of data [9,20] and informed consent [20,36,37]. In qualitative research, one challenge with ethics is anonymity of research participant, since both thick description and quotation from an interview might make it possible to recognize the participant [9]. The analyst should take care to assure anonymity. Ethical considerations, specific for HRA is not much discussed in literature, but it is a topic that should be further explored.

This should be reported in an HRA: The analyst should report how ethics were handled in the analysis and which ethical guidelines were followed. Some professional bodies might have their own ethical guidelines that are similar to the research guideline on ethics. The report should also discuss if there are some particular concerns with ethics.

4. Conclusion

In this paper, qualitative criteria for the qualitative part of HRA are presented. Both the International HRA empirical study [54] and the U.S. HRA empirical study [55] emphasize the importance of a good qualitative data analysis within HRA. Even if HRA results often are quantitative, most of the analysis performed is qualitative.

The qualitative part of an HRA is very much dependent on the analyst doing the analysis. A poor analyst might do a poor job with any HRA method. In this paper, how the analyst could do HRA in such a

way that it is possible for a reader or reviewer to judge the quality of the analysis is described. The criteria for qualitative analysis do not guarantee that good interpretation is done during the qualitative data analysis [20]. However, it increases the odds for high quality [20] and gives the reader and reviewer means of evaluating the interpretations made and the work presented.

Some HRA analysts might say that doing a proper qualitative analysis takes too much time and resources. However, I would argue that it is necessary to use quality criteria from qualitative research to improve the quality of the analysis; at least, the analysis should try to move closer to these criteria. HRA analysts should also be more trained in performing qualitative data collection and analysis, since this is an important part of the analysis.

References

- [1] Swain AD. Human reliability analysis: Need, status, trends and limitations. *Reliab Eng Syst Saf* 1990;29:301–13.
- [2] Kirwan B. *A guide to practical human reliability assessment*. Boca Raton, FL: CRC Press; 1994.
- [3] Boring RL, Blackman HS. The origins of the SPAR-H method's performance-shaping factor multipliers. Human factors and powerplants and HPRCT 13th annual meeting, IEEE8th IEEE; 2007. p. 177–84. <https://doi.org/10.1109/HFPP.2007.4413202>.
- [4] Gertman D, Blackman H, Byers J, Haney L, Smith C, Marble J. *The SPAR-H Method, NUREG/CR-6883*. Washington: DC: US Nuclear Regulatory Commission; 2005.
- [5] Bye A, Laumann K, Taylor C, Rasmussen M, Øie S, Van de Merwe K, et al. *The Petro-HRA Guideline*. Institute for Energy Technology, IFE/HR/E-2017/001. Halden, Norway. Downloaded from: Accessed 30 August 2017]. <https://www.ife.no/no/publications/2017/mto/the-petro-hra-guideline>; 2017 Accessed 30 August 2017].
- [6] Liao H, Forester J, Dang VN, Bye A, Lois E, Chang J. Lessons learned from the US HRA empirical study. *Probabilistic Safety Assessment and Management (PSAM 12)*. 2014 (2017) Downloaded from <https://pdfs.semanticscholar.org/db74/fc42608518635d11eab8a9c1001ecbffe1e.pdf> Accessed 30 August 2017.
- [7] Kirwan B, Ainsworth LK. *A guide to task analysis*. The task analysis working group. London, UK: Taylor & Francis; 1992.
- [8] Malterud K. Qualitative research: standards, challenges and guidelines. *Lancet* 2001;358:483–8.
- [9] Anastas JW. Quality in qualitative evaluation: Issues and possible answers. *Res Soc Work Pract* 2004;14:57–65.
- [10] Cepeda G, Martin D. A review of case studies publishing in *Management Decision* 2003–2004. *Manag Decis* 2005;43:851–76.
- [11] Levitt HM, Motulsky SL, Wertz FJ, Morrow SL, Ponterotto JG. Recommendations for designing and reviewing qualitative research in psychology: promoting methodological integrity. *Qual Psychol* 2017;4:2–22. doi:10.10377qap/0000082.
- [12] Jeanfreau SG, Leonard Jr J. Appraising qualitative research in health education: guidelines for public health educators. *Health Promot Pract* 2010;11:612–7. <https://doi.org/10.1177/1524839910363537>.
- [13] Elliott R, Fischer CT, Rennie DL. Evolving guidelines for publication of qualitative research studies in psychology and related fields. *Br J Clin Psychol* 1999;38:215–29.
- [14] Brantlinger E, Jimenez R, Klingner J, Pugach M, Richardson V. Qualitative studies in special education. *Except Child* 2005;71:195–207.
- [15] Meyrick J. What is good qualitative research? A first step towards a comprehensive approach to judging rigour/quality. *J Health Psychol* 2006;11:799–808. <https://doi.org/10.1177/1359105306066643>.
- [16] McWilliam RA. Author and reviewer guideline: reporting qualitative studies. *J Early Interv* 2000;23:77–80.
- [17] Onwuegbuzie AJ, Leech NL. Validity and qualitative research: An oxymoron? *Quality & Quantity* 2007;41:233–249.
- [18] Barusch A, Gringeri C, George M. Rigor in qualitative social work research: a review of strategies used in published articles. *Soc Work Res* 2011;35:11–9.
- [19] Yardley L. Demonstrating validity in qualitative psychology. In: Smith JA, editor. *Qualitative Psychology. A practical guide to research methods*. editor. Los Angeles, CA: Sage Publications; 2015. p. 257–72. 3rd ed.
- [20] Tracy SJ. Qualitative quality: Eight eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry* 2010;16:837–851, doi:10.1177/1077800410383121.
- [21] Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. *Qual Health Res* 2015;25:1212–22. <https://doi.org/10.1177/1049732315588501>.
- [22] Guba EG, Lincoln YS. Competing paradigms in qualitative research. In: Denzin NK, Lincoln YS, editors. *Handbook of qualitative research*. editors. Thousand Oaks, CA: Sage Publications; 1994. p. 105–17.
- [23] Guba EG, Lincoln YS. *Fourth generation evaluation*. Thousand Oaks, CA: Sage Publications; 1989.
- [24] Lincoln YS, Guba EG. *Naturalistic inquiry*. Thousand Oaks, CA: Sage Publications; 1985.
- [25] Rolfe G. Validity, trustworthiness and rigour: quality and the idea of qualitative research. *Methodol Issues Nurs Res* 2006;53:304–10.
- [26] Patton MQ. *Qualitative research & evaluation methods*. 4th ed. Thousand Oaks CA: Sage Publications; 2015.
- [27] Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101. <https://doi.org/10.1191/1478088706qp0630a>.
- [28] Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32 item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.
- [29] Morrow SL. Quality and trustworthiness in qualitative research in counseling psychology. *J Couns Psychol* 2005;52:250–60.
- [30] Yardley J. Dilemmas in qualitative health research. *Psychol Health* 2013;15:215–28. <https://doi.org/10.1080/0887044008400302>.
- [31] Angen MJ. Pearls, pith and provocation. Evaluating interpretive inquiry: reviewing the validity debate and opening the dialogue. *Qual Health Res* 2000;10:378–95.
- [32] Sobh R, Perry C. Research design and data analysis in realism research. *Eur J Mark* 2006;40:1194–209. <https://doi.org/10.1108/03090560610702777>.
- [33] Healy M., Perry C. Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative Market Research: An International Journal* 2000;3:118–126 doi:10.1108/13522750010333861.
- [34] Ali AM, Yusuf H. Quality in qualitative studies: the case of validity, reliability and generalizability. *Issues Soc Environ Account* 2011;5:25–64.
- [35] Beverland M, Lindgren A. What makes a good case study? A positivist review of qualitative case research in industrial marketing management. *Ind Mark Manag* 2010;39:56–63. <https://doi.org/10.1016/j.indmarman.2008.09.005>.
- [36] Drisko JW. Strengthening qualitative studies and reports. *J Soc Work Educ* 1997;33:185–97. <https://doi.org/10.1080/10437797.1997.10778862>.
- [37] O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014;89:1245–51. <https://doi.org/10.1097/ACM.0000000000000388>.
- [38] Cohen DJ, Crabtree BF. Evaluative criteria for qualitative research in health care: controversies and recommendations. *Ann Fam Med* 2008;6:331–9.
- [39] Long T, Johnson M. Rigour, reliability and validity in qualitative research. *Clin Effect Nurs* 2000;4:30–7.
- [40] Riege AM. Validity and reliability tests in case study research: a literature review with ‘hands-on’ applications for each research phase. *Qual Mark Res* 2003;6:75–86. <https://doi.org/10.1108/13522750310470055>.
- [41] Zitomer MR, Goodwin D. Gauging the quality of qualitative research in adapted physical activity. *Adapt Phys Act Q* 2014;31:193–218. <https://doi.org/10.1123/apaq.2013-0084>.
- [42] Pilnick A, Swift JA. Qualitative research in nutrition and dietetics: assessing quality. *J Hum Nutr Diet* 2010;24:209–14. <https://doi.org/10.1111/j.1365-277X.2010.01120.x>.
- [43] Amankwa L. Creating protocols for trustworthiness in qualitative research. *J Cult Divers* 2015;23:121–7.
- [44] Ravenek MJ, Rudman DL. Bridging conceptions of quality in moments of qualitative research. *Int J Qual Methods* 2013;12:436–56.
- [45] Williams EN, Morrow SL. Achieving trustworthiness in qualitative research: a pan-paradigmatic perspective. *Psychother Res* 2009;19:576–82. <https://doi.org/10.1080/10503300802702113>.
- [46] Lewis J. Redefining qualitative methods: believability in the fifth moment. *Int J Qual Methods* 2009;8:1–14.
- [47] Kitto SC, Chesters J, Grbich C. Quality in qualitative research Criteria for authors and assessors in the submission and assessment of qualitative research articles for the *Medical Journal of Australia* *Med J Aust* 2008;188:243–6.
- [48] Kuper A, Lingard L, Levinson W. Critically appraising qualitative research. *BMJ* 2008;337:687–92. <https://doi.org/10.1136/bmj.a1035>.
- [49] Kline WB. Developing and submitting credible qualitative manuscripts. *Couns Educ Superv* 2008;47:210–7.
- [50] Ryan-Nicholls KD, Will CI. Rigour in qualitative research: mechanisms for control. *Nurse Res* 2009;16:70–85.
- [51] Brooks J, King N. Qualitative psychology in the real world: the utility of template analysis. *British psychological society annual conference, 18–20 April, 2012*. 2012 Downloaded from <http://eprints.hud.ac.uk/id/eprint/13656/1/Brooks> Accessed 30. August 2017.
- [52] Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. *Qual Res Psychol* 2015;12:202–22. <https://doi.org/10.1080/14780887.2014.955224>.
- [53] King N, Cassell C. *Doing template analysis*. In: Symon G, editor. *Qualitative organizational research*. editors. London: UK: Sage Publishing; 2012. p. 426–50.
- [54] US Nuclear Regulatory Commission. The international HRA empirical study: lessons learned from comparing HRA methods predictions to HAMMLAB simulator data, NUREG-2127. US Nuclear Regulatory Commission. 2014. p. 1–157.
- [55] US Nuclear Regulatory Commission. The U.S. HRA empirical study – assessment of HRA method predictions against operating crew performance on a U.S. Nuclear Power Plant Simulator (NUREG-2156). US Nuclear Regulatory Commission. 2016. p. 1–461.