Suggestions for improvements to the definitions of SPAR-H performance shaping factors, to the definitions of the levels, and suggestions for changes in the multipliers.

Karin Laumann[a] and Martin Rasmussen[a]
[a] Norwegian University of Science and Technology, Trondheim, Norway

Abstract: In this paper the definitions and the content of six of SPAR-H performance shaping factors are discussed. The six factors discussed are "Available time", "Stress/Stressors", "Experience/Training", "Procedures", "Fitness for Duty" and "Work Processes". The discussion is based on a literature study on performance shaping factors, on interviews with consultants that have done SPAR-H analysis in the petroleum industry and from reading Human Reliability Analysis reports where SPAR-H have been used. The conclusions in this paper are: 1) New description of SPAR-H PSFs should be developed where the descriptions of each PSF do not overlap so much. 2) The guidelines should also give more advice to help the analyst to select PSFs levels when multiple PSFs might have a positive or negative impact on error probabilities. 3) New multipliers should be developed from an expert judgment, which is based on a review of the existing literature on PSFs, and with knowledge of the work in control room today,

Keywords: HRA, PSF, SPAR-H, Human performance, Human Error Probabilities.

1. INTRODUCTION

The purpose of this paper is to investigate the definitions of the performance shaping factors (PSFs), the definitions of the levels of the PSFs, and the multipliers in the Human Reliability Analysis (HRA) SPAR-H. The SPAR-H method has been described in several guidelines (1,2,3).

A PSF is a factor that has a negative or positive impact on human performance. SPAR-H includes eight PSFs. This paper will discuss the definitions, levels and multipliers of six of these factors; "Available time", "Stress/Stressors", "Experience/Training", "Procedures", "Fitness for Duty", and "Work Processes". We have discussed the "Ergonomics/HMI" PSF in another paper submitted to PSAM12 (4). We have discussed the definition, levels and multipliers of the "Complexity" PSF in another paper (5).

This paper is based on:
1) An investigation of the definitions, levels and multipliers in the SPAR-H guidelines and where they originated from.
2) A literature review on the performance shaping factors and how they impact human performance.
3) Interviews with consultants who have performed HRA analysis with SPAR-H in the petroleum industry.
4) Investigation of some of the HRA reports from the petroleum industry where SPAR-H has been used.

This paper presents work in progress. It presents the suggestions that we have so far for changes to the quantification method in SPAR-H. The paper is meant to be a basis for discussions about these changes. It should not be looked at as our final conclusions about changes to the SPAR-H quantification method. Also, more work is needed to develop new descriptions of the SPAR-H PSFs, levels and multipliers. We will continue with this work during 2014 and our aim is to end up with an improved SPAR-H guideline for use in the petroleum industry.

2. SUGGESTED CHANGES TO SPAR-H PERFORMANCE SHAPING FACTORS, LEVELS AND MULTIPLIERS

2.1. Available Time

The definition of "Available time" in NUREG/CR-6883; The SPAR-H Human Reliability Analysis Method (1) is: "Available time refers to the amount of time that an operator or a crew has to diagnose and act upon an abnormal event. A shortage of time can affect the operator's ability to think clearly and consider alternatives. It may also affect the operator's ability to perform. Multipliers differ somewhat, depending on whether the activity is a diagnosis activity or an action."

Something that is peculiar about the SPAR-H definition of the "Available time" PSF is that time pressure is not defined as available time, but as part of the "Stress/Stressors" PSF. We think that time pressure is the key element of "Available time" and that it should therefore be included in the "Available time" PSF. In one of the SPAR-H guidelines (3, page 12) it is also stated about "Available time" that; "It can also refer to perceived time pressure: if operators feel there is insufficient time, or if they perceive a need to hurry, regardless of the objective time available to diagnose a problem or complete a task, performance decrement can occur." So there is also some discrepancy within the SPAR-H guidelines on where time pressure should be included.

Boring (6) found that analyst evaluations of the "Available time" PSF and the "Stress/Stressors" PSF in SPAR-H correlated .67 for diagnosis and .50 for actions. An explanation for the high correlations between these two factors could be the overlapping definitions of content in these two factors, since a shortage of time available usually causes time pressure. The confusion in what to attribute where in SPAR-H is problematic both in terms of potential double counting and reduced inter-analyst reliability.

The SPAR-H guidelines give the analyst very little advice on how to analyse how long time a task will take for the operators' to perform. This could be challenging for an inexperienced analyst. Advice about how to analyse how long time it takes for an operator to perform a task, could be found in existing human factors guidelines and from interviews with experienced human factors experts. These advices should include consideration not only on how long time it would take to do a task, but also how the context could affect how much time an operator or a crew would use (such as the scenario, distractions, time to read the procedures, crew characteristics as for example fast and aggressive responders or slow and methodical, and how the crew is trained).

The levels in SPAR-H for "Available time" for diagnosis (At-power Condition) are:
Inadequate time - If the operators cannot diagnose the problem in the amount of time available, no matter what s/he does, then failure is certain. Multiplier: P(failure) = 1.0
Barely adequate time – 2/3 the average time required to diagnose the problem is available. Multiplier: 10
Nominal time – on average there is sufficient time to diagnose the problem. Multiplier: 1.
Extra time – time available is between one to two times greater than the nominal time required, and also greater than 30 minutes. Multiplier: 0.1.
Expansive time – time available is greater than two times the nominal time required and is also greater than a minimum time of 30 minutes; there is an inordinate amount of time (a day or more) to diagnose the problem. Multiplier: 0.01
Insufficient information – if you do not have sufficient information to choose among the other alternatives, assign this PSF level. Multiplier: 1.

Boring and Blackman (7) describe that the multipliers for "Available time" in SPAR-H "align with THERP Table 20-1." "Item 2" in this Table in THERP (8) was however not included. "Item 2" is less

than 10 minutes after annunciator. This item has a Human Error Probability (HEP) of .50. It could be argued that the "Barely adequate time" seems more to fit with 10 minutes after an annunciator and a multiplier of 50 than with 20 minutes after an annunciator and a multiplier of 10. Boring and Blackman (7) gives no argument to why item 2 i table 20-1 in THERP was left out.

Reer and Sträter (9) have presented a table of time reliability curves from different data-based methods. They have presented comparisons of HRA methods (THERP nominal curve, EDF curves and HRC) on time available to diagnose a task and expected human failure rates. All of these methods have a HEP of 0.5 for the first minutes after an event has started. However, the methods differ in how many minutes available after an event the HEP 0.5 level should be estimated. Also for the HEP of 0.1, the methods to some degree differ in how many minutes after the annunciator the operators have available for the 0.1 HEP should be used.

The multiplier for the level "expansive time" seems also to be very optimistic in SPAR-H with a multiplier 0.01.

From comparison with THERP (8) and Reer and Sträter's study (9), the levels and multipliers in SPAR-H could be:

Operator(s) do not have enough time/inadequate time. Multiplier: P(failure) = 1.0.
Extreme time pressure (Barely enough time to do the task). Multiplier: 50.
High/moderate time pressure. Multiplier: 10.
Nominal time/time constrain do not affect operators performance neither negative nor positive. Multiplier: 1.
Extra time, there is good time/extra time to perform the task. Multiplier: 0.1.

More work is needed to define the different levels in time (minutes) required to be at that level.

2.2. Stress/Stressors

The definition of the "Stress/Stressors" PSF in NUREG/CR-6883; The SPAR-H Human Reliability Analysis Method (1) is: "Stress (and level of arousal) have been broadly defined and used to describe negative as well as positive motivating forces of human performance. Stress as used in SPAR-H refers to the level of undesirable conditions and circumstances that impede the operator from easily completing a task. Stress can include mental stress, excessive workload, or physical stress (such as that imposed by difficult environmental factors). It includes aspect of narrowed attentional field or muscular tension, and can include general apprehension or nervousness associated with the importance of an event. Environmental factors often referred to as stressors, such as excessive heat, noise, poor ventilation, or radiation, can include stress in a person and affect the operator's mental and physical performance. It is important to note that the effect of stress on performance is curvilinear –some small amount of stress can enhance performance, and should be considered nominal, while high and extreme level of stress will negatively affect human performance. "

Stress has a very large variety of definitions including the external stimuli that causes stress, the internal stress, the stress reaction, or the outcome of stress. Salas, Driskell and Hughes (10) have summed up the different definitions and meanings of the stress concept and they have developed a four step model of stress and performance. The four steps are:

1) An environmental stimulus becomes salient.
2) It requires a positive or negative valiance though the appraisal process.
3) It leads to the formation of performance expectation.
4) And these in turn determine the number of physiological, cognitive, emotional and social consequences.

This model sums up the literature on stress and it is consistent with the most influential stress theories. From this model one can see that it is difficult to separate the "Stress/Stressors" PSF, as it is defined in SPAR-H, from the other PSFs in SPAR-H. Every other PSF could be thought of as an environmental stressor. Time pressure for example could be included in the "Available time" PSF, noise, temperature and difficult environmental factors seems to belong to the "Ergonomics/HMI" PSF. Noise and temperature does however appear to only have a slight effect on performance and should perhaps not be included in the Petro-HRA method (4). It could be argued that the "Stress/Stressors" PSF is not a required PSF in SPAR-H because all the other PSFs are the environmental "Stress/Stressors." However there is one aspect of the "Stress/Stressors" PSF in SPAR-H that is not covered by the other PSFs and that is the concept of "threat stress". Driskell and Salas (10, page 23) defines threat stress as: "the anticipation of fear of physical or psychological harm […] thus a threat provoking situation is one in which dangerous and novel environmental events pose the potential for pain and discomfort."

Because of human and ethical concerns it is very difficult to do experimental studies on threat stress. Swain and Guttmann (8) refer to some "old" studies that they used to define the HEP for stress in THERP which could be considered to estimate the effect on performance from threat stress. The effects of threat stress on performance have probably not changed much since the time these studies where done. There are also some other studies that have been done that could be used to estimate the effect of threat stress on human performance. Driskell et al. (11) for example have performed a meta-analysis on the effect of threat of electrical shock on performance accuracy and subjective experience of threat. There are also some studies on performance under different from of stress such as for example jumpmaster training (12).

New levels for "Threat Stress" in SPAR-H could be:
Extreme "Threat Stress"-the operators' life is in immediate danger. For example with a fire in the control room. Multiplier P (failure) = 1.
High "Threat Stress" – the operators think that their own and others life are in danger.
Moderate "Threat Stress" – the operator's fear that doing mistakes have a large negative effect on self-esteem or professional status.
Nominal threat stress – "Threat Stress" does not seem to affect performance.

New multipliers for each of the levels of "Threat Stress" should be defined by expert judgment which is based on the findings in the available literature on "Threat Stress" impact on human performance.

2.3. Experience/Training

The definition of the "Experience/Training" PSF in NUREG/CR-6883; The SPAR-H Human Reliability Analysis Method (1) is: "This PSF refers to the experience and training of the operator(s) involved in the task. Included in this consideration are years of experience of the individual or crew, and whether or not the operator/crew has been trained on the type of accident, the amount of time passed since training and the systems involved in the task and scenario. Another consideration is whether or not the scenario is novel or unique (i.e., whether or not the crew or individual has been involved in a similar scenario, in either a training or an operational setting). Specific examples where training might be deficient are guidance for bypassing engineered safety functions, guidance for monitoring reactor conditions during reactivity changes, and guidance for monitoring plant operation during apparently normal, stable conditions for the purpose of promoting the early detection of abnormalities."

The definitions of the levels in NUREG/CR-6883 are:
Low – less than 6 months of experience and/or training. This level of experience/training does not provide the level of knowledge and deep understanding required to adequately perform the required tasks; does not provide adequate practice in those tasks; or does not expose individuals to various abnormal conditions.

Nominal – more than 6 months experience and/or training. This level of experience/training provides an adequate amount of formal schooling and instruction to ensure that individuals are proficient in day-to-day operations and have been exposed to abnormal conditions.
High – extensive experience; a demonstrated master. This level of experience/training provides operators with extensive knowledge and practice in a wide range of potential scenarios. Good training makes operators well prepared for possible situations.

There have been several meta-analyses on the relationship between work experience and general work performance which has found low to medium correlation between work experience and general work performance (13,14). However, work experience can be defined in many different ways. Quinones, Ford and Teachout (15) have done both a conceptual review of the meanings of work experience and a meta-analysis to test different types of job experience on performance. Quinones et al. (15) sum up their findings in their meta-analysis (page 904): "The result of the meta-analysis revealed that the relationship between work experience and job performance was positive regardless of the work experience measurement used. The relationship was stronger when hard performance measures such as work samples were used as compared to soft performance measures such as supervisory ratings. The meta-analyses results also revealed some variations in the relationship between work experience and job performance as a function of measurement mode. The strongest relationship occurred between amount of experience and performance. Time and type measures showed the weakest relationship. Finally, variability in the relationship between experience and performance as a function of level of specificity was found. Task-level experience had the strongest relationship with performance whereas organizational level showed the weakest. These results are consistent with expectations and suggest that various measures of work experience capture different aspects of job-relevant experience. Amount and task-level measures are perhaps better measures of what individuals actually do on the job. Time-based measures are likely to be poor indicators of actual experience. Similarly task-level measures may capture more specific experiences than do job or organizational measures."

In SPAR-H the PSF levels on the "Experience/Training" PSF is defined by "how long time the operators have been in the job." A better definition of the PSF levels seems to be an amount and task related definition. With such a definition the analyst will look at how many times or how often the operators have either performed or trained on the scenario or a task in question. In Quinones et al. study, when work experience was defined as amount mean corrected correlation between experience and training was .43 and when experience was measured at the task level the mean corrected correlation was .41.

The most adverse PSF level for "Experience/Training" in SPAR-H has a multiplier of 10 (for diagnosis). This multiplier seems very low compared to for example the multiplier in HEART for error producing condition with a similar content as "Experience/Training" in SPAR-H. It is also a factor that is considered as one of the most important factors that an organization can influence to increase human reliability (16). We would therefore argue that we would expect the "Experience/Training" PSF to have one of the highest multipliers in SPAR-H. Our suggestions for levels and multipliers are:

There is a mismatch between experience/training and the task or scenario that is analyzed. Multiplier: P(failure) = 1.0
There is no training on scenario and task. Multiplier: 50
The operators have some experience/training with the scenario/task(s) but the experience/training is considered low. Multiplier: 10
Nominal, the operators have experience/training on the scenario/task(s). Multiplier: 1.
Good experience/training with the task and scenario. The operators have extensive training on the task and scenario. Multiplier: 0.5.

2.4. Procedures

The definition of the "Procedures" PSF in NUREG/CR-6883; The SPAR-H Human Reliability Analysis Method (1) is: "This PSF refers to the existence and use of formal operating procedures for the tasks under consideration. Common problems seen in event investigations for procedures include situations where procedures give wrong or inadequate information regarding a particular control sequence. Another common problem is the ambiguity of steps. PSF levels differ somewhat, depending on whether the activity is a diagnosis activity or an action. In situations where multiple transitions between procedures are required to support a task or group of tasks, SPAR-H suggests that the analyst adjust the PSF for complexity accordingly. If the procedures themselves are problematic, i.e., inadequate, then, the HRA analyst should assess the procedures and determine whether they should be assigned an "inadequate" or "poor" rating".

The levels and multipliers for "Procedures" for diagnosis are:
Not available – the procedure needed for å particular task or tasks in the event is not available. Multiplier: 50.
Incomplete – information is needed that is not contained in the procedure or procedure sections; section or task instructions (or other needed information) are absent. Multiplier: 20.
Available, put poor – a procedure is available but it is difficult to use because of factors such as formatting problems, ambiguity, or such a lack in consistency that it impedes performance. Multiplier: 5.
Nominal – procedures are available and enhance performance, multiplier: 1.
Diagnostic/symptom oriented – diagnostic procedures assist the operator/crew in correctly diagnosing the event. Symptom-oriented procedures (sometimes called function-oriented procedures) provide the means to maintain critical safety functions. These procedures allow operators to maintain the plant in a safe condition, without the need to diagnose exactly what the event is, and what needs to be done to mitigate the event. There will be no catastrophic result (i.e., fuel damage) if critical safety functions are maintained. Therefore, if either diagnostic procedures (which assist in determining probable cause) or symptom oriented procedures (which maintain critical safety functions) are used, there is less probability that human error will lead to a negative consequence. This being said, if the symptom-based procedure is found to be inaccurate or awkwardly constructed, then the procedures PSF should be negatively rated. Multiplier: 0.5
Insufficient information – if you do not have sufficient information to choose among the other alternative assign this PSF level: 1.

We suggest that the availability/quality of procedures and the use of the procedures are evaluated in the same PSF. It seems illogical to split these two aspects of procedures as done in the SPAR-H guideline. According to the SPAR-H guidelines (1) the availability and quality of the procedures are evaluated in the "Procedures" PSF, while the use of the procedures is evaluated in the "Work Processes" PSF. It seems for example illogical to get a lower HEP for having diagnostic/symptom oriented procedures if they are not also used in a good way. It also seems illogical to have a higher multiplier for procedures "not available" (multiplier= 50, from the "not available" level of the "Procedures" PSF) than for procedures "not used" (multiplier = 5, from the "Poor" level in the "Work Processes" PSF). It could be argued that in these two situations, the "Procedures" PSF should have the same effect on performance. If there are no procedures available, the operators cannot use them, leading to a difficult situation for the analyst who has to decide whether to choose a multiplier from the level "not available" from "Procedures" PSF or the multiplier for "Poor" in the "Work Process" PSF, or both of them.

When, "how the procedure is used," is included in the "Work Processes" PSF it becomes more difficult for an analyst to decide which part of the PSF belongs to the "Procedure" PSF and which belong to the "Work Processes" PSF, especially since these two aspects is highly correlated. When these two aspects of procedures are defined in two separate PSFs it creates an unnecessary overlap between the content of these two PSFs and it reduces the clarity of the method and different analysts will attribute different PSFs when analyzing the same situations. In Boring's (6) study the correlation

between the "Procedures" PSF and the "Work Processes" PSF was .36 for diagnosis and .35 for action. These correlations are probably partly reflecting the shared defined content between these two factors.

It seems also illogical that the SPAR-H guidelines advice the analyst to: "In situations where multiple transitions between procedures are required to support a task or group of tasks, SPAR-H suggests that the analyst adjust the PSF for complexity accordingly. If the procedures themselves are problematic, i.e., inadequate, then, the HRA analyst should assess the procedures and determine whether they should be assigned an "inadequate" or "poor" rating." Multiple transitions between procedures seem to be a problem with the procedures and should be evaluated within the "Procedures" PSF and not within the "Complexity PSF". Advices like this in SPAR-H increase the ambiguity of the guideline on which PSF to attribute and this has the potential to reduce inter-rater reliability.

The levels for the "Procedure" PSF could be:
Misleading Procedures: The procedures are strongly misleading for the scenario or task in question and they are used by the operators. Multiplier: P(failure) = 1.
No procedures available or the procedures are not used. There are no procedures or the operators do not use the procedures. Multiplier: 50.
Poor procedures: There are procedures available but there are human factors issues with the procedure or the use of the procedure Multiplier: 20.
Nominal procedures: There are good procedures with no human factors issues and they are used by the operators. Multiplier: 1.

We recommend that the "Procedures" PSF should not include a "good" level because having good procedures with no human factors issues, and that the procedures are used," should be considered as the nominal in a high risk industry.

2.5. Fitness for Duty

The definition of the "Fitness for Duty" PSF in NUREG/CR-6883; The SPAR-H Human Reliability Analysis Method (1) is: "Fitness for duty refers to whether or not the individual performing the task is physically and mentally fit to perform the task at the time. Things that may affect fitness include fatigue, sickness, drug use (legal or illegal), overconfidence, personal problems and distractions. Fitness for duty includes factors associated with individuals, but not related to training, experience, or stress."

Whaley, Kelly, Boring and Galyean (2) say that issues like impairment to drugs or alcohol, distraction due to family issues, whether a person is physically or mentally capable of performing a task or boredom are rarely documented in event reports and that the most common fitness for duty issue cited is fatigue. The "Fitness for Duty" PSF is seldom selected in a prospective analysis. It is very rarely that a safety critical organization has operators that on a general basis are unfit for duty. The aspects outside of fatigue included in this PSF are so unlikely to affect performance that they should not be included into a method like SPAR-H. Fatigue, on the other hand, is a factor that often will influence HEP in a significant manner and a topic that is extensively studied and there exists several meta-studies on the effect of fatigue on performance. To reduce confusion and the potential negative associations companies have towards evaluating the fitness for duty in their employees, we suggest that this PSF should be named "Fatigue".

Williamson et al. (17) have defined fatigued as a "biological drive for recuperative rest." They say that fatigue may include sleepiness, but also mental, physical and muscular fatigue depending of the nature of its cause. Williamson et al. (17) have developed a model of the relationship between fatigue and safety. In the model there are three factors that cause fatigue and sleepiness; time of day, time awake and task related factors such as a task that either demands sustained attention or monotony over a longer time period.

Folkard and Lombardi (18) performed a meta-analysis on fatigue by comparing accident rates on day-, afternoon- and nightshifts. In this meta-analysis they have only included studies that have controlled for other factors than fatigue that can vary between different shifts (such as the number of people on duty, tasks that are done, reporting practices and experience/competence). Folkard and Lombardi (18) concluded from their results that incidents increased in an approximately linear fashion; from the morning to the afternoon shift the increase was 15.2 percent and from the morning to the night shift the increase were 27.9 percent. This meta-analysis also includes a comparison of incidents across successive hours at work and over successive shift days. Studies like this should be included into an expert evaluation to define PSF levels and multipliers for the "Fatigue" PSF.

The levels for "Fatigue" could be:
Extreme fatigue: the fatigue level is so high that the operators cannot perform the task. Multiplier: P(failure) = 1.0.
High fatigue: A very long lasting demanding task or a very long lasting monotonous task. Multiplier: 5
Moderate fatigue: Task on a night shift. Multiplier: 1.3.
Nominal: Fatigue does not affect the task. Multiplier: 1.

2.6 Work Processes

The definition of the "Work processes" PSF in NUREG/CR-6883; The SPAR-H Human Reliability Analysis Method (1) is: "Work processes refer to aspects of doing work, including inter-organizational, safety culture, work planning, communication, and management, support and policies. How work is planned, communicated and executed can affect individual and crew performance. If planning and communication are poor, then individuals may not fully understand the work requirements. Work processes include consideration of coordination, command and control. Work processes also include any management, organizational or supervisory factors that may affect performance. Examples seen in event investigations are problems due to information not being communicated during shift turnover, as well as communication with maintenance crews and auxiliary operators. Measures could include the amount of rework, risk worth of items in utility corrective action program backlog, enforcement action, turnover, performance efficiencies, etc.

The shift supervisor also plays a major role in work processes. Instances where the shift supervisor gets too involved in the specifics of the event – in contrast to maintaining a position of leadership in the control room – would indicate a breakdown in work processes.

Conditions with effects adverse to quality are also included in the work practice category, as are problems associated with a safety-conscious work environment. This includes retaliation by management against allegations as it pertains to the failure event under investigation. For example, the analyst must decide whether utility management actions against maintenance staff have any bearing on a particular control room action or maintenance action under evaluation. If the analyst believes there is such evidence, then the appropriate negative level for work practice is assigned.

Additionally, any evidence obtained during the review of an operation event indicating inter-group conflict and decisiveness (e.g., between engineering and operations) or an uncoordinated approach to safety, is evaluated in SPAR-H as a work process problems. Schisms between operators and management are also considered work process problems.

SPAR-H does directly acknowledge potential problems between the regulator and licensee as it may affect operator and crew performance. It is assumed that problems in communication or adherence to enforcement actions or notices are indicative of work process problems.

Finally, inadequacies in the utility corrective action program (CAP), such as failure to prioritize, failure to implement, failure to respond to industry notice, or failure to perform root cause as required by regulation, is considered in SPAR-H as a work process variable. Because there are so many

potential areas of concern with the work process category that can be assigned to a potential PSF level, the analyst is directed to provide as much information as possible in the worksheet space provided, listing the reasons for assigning a particular work process PSF level."

The "Work Processes" PSF in SPAR-H is so poorly defined that it is difficult to say which concepts that should be evaluated under this PSFs. It is also difficult to know which concepts one should search for in a literature study to explore how work processes impact performance.

For us, the "Work Processes" PSF in SPAR-H seems to consist of: Safety culture, teamwork (teamwork also includes team leadership) and use of procedures and documentation. Elements in the definition of the "Work Processes" PSF such as planning /scheduling and administrative control seem to be, tasks in itself, and not PSFs. We have earlier discussed that "use of procedures" should be moved to the "Procedures" PSF.

After reviewing the literature on safety culture and teamwork, we decided that the PSF "Work Processes" should be split into two PSFs, one related to safety culture and one related to teamwork. First the conclusions related to safety culture from this review are presented then the conclusions from the review related to teamwork are presented.

Guldenmund (19. page 1466) stated: "Safety culture has become a term used by people all around the globe to explain everything relating to safety failures that cannot be explained in another way. That the concept is fuzzy does not seem to matter much; however, this fuzziness is both its strength and its weakness. Indeed, (groups of) people sometimes seem to perform in dark, mysterious ways and when grouping for an explanation fuzzy concept such as safety culture is highly attractive. A similar, initial attractiveness is inherent in the concept of organizational culture."

Safety culture seems to be a concept that included everything a high reliable organisation does to be safe. This makes the concept extremely unclear and it also overlaps with other PSFs in SPAR-H (especially procedures and experience/training). The concept is so broad that it is a question if it is meaningful in a method like SPAR-H which aims at using very specific factors to predict the likelihood than an operator or a crew of operators fails to do a task. To assure inter-rater reliability and validity of the method it should be clear what the PSF includes and not includes. Also how a PSF is expected to affect operators' performance of an activity has to be described and understood. If we do not know how a PSF affects performance it is impossible so decide on a multiplier.

An analysis of safety culture/safety climate and an HRA analysis are very different. These types of analyses come from different research areas with different purposes. A SPAR-H analysis is a much more specific analysis where the likelihood of human failure in a post-initiation activity is analysed. HRA is linked to psychological experimental studies on factors that affect task performance. A safety culture analysis, investigates safety of a much broader organizational context, or the total organization. It included much more distal organizational concepts and is not easy to describe how they affect operators' performance.

The concept safety culture is too broad to be well defined in SPAR-H. This concept should be taken out and replaced with organizational factors that are well defined and which have shown to have a medium to large significant contribution to operator's tasks for post initiator events.

This literature study did not identify any study which has tested how the safety culture conditions contribute to human errors in a post-accident scenario. One factor that has been mentioned in accident investigation such as Deep Water Horizon accident (20) as an important factor that affected operators performance is; A goal conflict between safety and production. The term "Work Processes" in SPAR-H could be called "Attitudes towards safety" and then included factors such as;

- Prioritising safety over production (and other goals) when that is appropriate.
- Attitudes to work and work conduct.

During the last ten to twenty years there has been an extensive research on teamwork. Salas, Sims and Burke (21) presented a literature review to define teamwork and which factors or dimensions that should be included in teamwork. This literature review was selected because it is comprehensive, it is relatively new and it describes behavioural markers for the teamwork activity. In Salas et al.'s (21) literature review, a thematic analysis of the variables most commonly discussed and having the greatest effect on team performance, was included. Salas et al. (21, page 558-562) defined a team: "as two or more individuals with specified roles interacting adaptively, interdependently, and dynamic toward a common and valued goal." Teamwork is defined as: "A set of interrelated thoughts, and feeling of team members that are needed to function as a team and that combine to facilitate coordinated, adaptive performance and task objective resulting in value-added outcomes."

Salas et al. (21) found five core components (team leadership, mutual performance modelling, backup behaviour, adaptively and team orienting) and three coordinating mechanisms (shared mental models, achievement of mutual trust and close-loop communication). In the literature review Salas et al. (20) describe these factors and present research on how these factors are related to team and task performance. The Salas et al. (21) review could be used to describe a teamwork PSF in SPAR-H.

Several meta-studies have tested the effect of team work factors on performance. One example is presented briefly: De Dreu, Carsten, and Weingart (22) have done a meta–analysis on the associations between relationship conflict, task conflict, team performance, and team member satisfaction. The results showed a strong and negative correlation between relationship conflict, team performance, and team member satisfaction. Conflict had stronger negative relations with team performance in highly complex (decision making, project, mixed) than in less complex (production) tasks. Task conflict was less negatively related to team performance when task conflict and relationship conflict were weakly, rather than strongly, correlated. The large amount of research on team work and performance could be used to define the multipliers for a teamwork PSF in SPAR-H

To sum up we suggest splitting the "Work Processes" PSF into these two factors:

1) Attitudes to safety and work conduct.
2) Teamwork.

The levels for attitudes to safety and work conduct could be:
Attitudes to safety and work conduct are extremely poor. Multiplier: P (Failure) = 1.0.
Attitudes so safety and work conduct are poor. Multiplier: 50.
Attitudes to safety and work conduct are to some degree poor. Multiplier: 10.
Attitudes to safety and work conduct do not affect performance neither negative nor positive. Multiplier = 1.
Attitudes to safety and work conduct are good. Multiplier: 0.5.

The level of team work could be:
Teamwork is needed for this task and teamwork is extremely poor. P (Failure) = 1.0.
Teamwork is needed for the task and teamwork processes is poor. Multiplier: 50.
Teamwork in needed for this task and some of the teamwork processes is poor. Multiplier: 10.
Teamwork is not needed or teamwork has neither a negative/nor positive effect on the task. Multiplier = 1.
Teamwork will have a positive effect on the task and teamwork is good. Multiplier: 0.5.

3. CONCLUSION

The conclusion from this paper is that the SPAR-H guidelines should be revised. There are several suggestions for how the guidelines could be improved:

1) The PSF descriptions should be improved, clearly defining the areas included in each PSF. This

should increase the inter-rater reliability, and reduce the overlap and correlations between the PSFs. This is a general issue in SPAR-H, but one that is particularly evident in the "Work Processes" PSF. It is difficult to obtain inter-rater reliability when it is unclear in which PSF a particular issue should be attributed in. Some issues could be attributed in more than one PSF leading to analysts ending up with very different HEPs due to differences in levels and multipliers in the SPAR-H PSFs.

2) The guidelines should give more advice to help the analyst to select PSF levels when multiple PSFs might have a positive or negative impact on the HEP.

3) Boring and Blackman (7) have described that most of the multipliers in SPAR-H originated from the THERP tables. THERP is over 30 years old. During these 30 years there has been an extensive amount of research on performance shaping factors. Also the structure of the work in a control room might have changes during these years. New multipliers should be developed from expert judgment, which is based on a review of the existing literature on PSFs, and from knowledge of the work in control room today.

Acknowledgements

This work has been carried out as part of The Research Council of Norway project number 220824/E30 "Analysis of human actions as barriers in major accidents in the petroleum industry, applicability of human reliability analysis methods (Petro-HRA)". Financial and other support from The Research Council of Norway, Statoil ASA and DNV are gratefully acknowledged. The authors will also thank the Office of Naval Research USA for travel funds for the PSAM 12 conference. This paper represents the opinion of the authors, and does not necessarily reflect any position or policy of the above mentioned organizations.

References

[1] D. Gertman, H. Blackman, J. Marble, J. Byers and C. Smith. "*The SPAR-H Human Reliability Analysis Method*," NUREG/CR-6883. INL/EXT-05-00509, (2005).
[2] A. M. Whaley, D. L. Kelly, R. L. Boring and W. J. Galyean. "*The SPAR-H Step-by-Step Guidance*." INL/EXT-10-18533. Rev 2, (2011).
[3] R. L. Boring, A. M. Whaley, T. Q. Train, P. H. McCabe, L. G. Blackwood and R. F. Buell. "*Guidance on performance shaping factor assignments in SPAR-H*." INL/EXT-06-11969, (2006).
[4] M. Rasmussen and K. Laumann. "*The suitability of the SPAR-H "Ergonomics/HMI" PSF in a computerized control room in the petroleum industry*." Paper submitted to PSAM12, (2014).
[5] M. Rasmussen, M. I. Standal and K. Laumann. "*Task complexity as a performance shaping factor: A review and recommendations in SPAR-H adaption*." Journal of Risk Research. (In review).
[6] R. Boring. " *How Many Performance Shaping Factors are Necessary for Human Reliability Analysis?"* No. INL/CON-10-18620. Idaho National Laboratory (INL), (2010).
[7] R. L. Boring and H. S. Blackman, "*The origins of the SPAR-H method's performance shaping factor multipliers*." In Human Factors and Power Plants and HPRCT 13th Annual Meeting, 2007 IEEE 8th (pp. 177-184), IEEE (2007).
[8] D.A. Swain and H. E. Guttmann. " *Handbook of Human reliability analysis with emphasis on nuclear power plant application*." NUREG/CR-1278. Washington, D.C., USA, (1983).
[9] B. Reer and O. Sträter. "*Evaluation of new developments in cognitive error modeling and quantification: time reliability correlation."* In Probabilistic Safety Assessment and Management'96 pp. 645-650. Springer London, (1996).
[10] E. Salas, J. E. Driskell and S. Hughes. "*Introduction: The study of stress and human performance*," In J. E. Driskell, and E. Salas, (Eds.). Stress and human performance. Psychology Press, (2013).
[11] J. E. Driskell, B. Mullen, C. Johnson, S. Hughes, and C. L. Batchelor. "*Development of quantitative specifications for simulating the stress environment*." ARMSTRONG LAB WRIGHT-PATTERSON AFB OH, (1992).

[12] W. P. Burke. "*Development of predictors of performance under stress in Jumpmaster training"* (No. ARI-RR-1352). ARMY RESEARCH INST FOR THE BEHAVIORAL AND SOCIAL SCIENCES ALEXANDRIA VA, (1980).
[13] J. E. Hunter and R. F. Hunter. "*Validity and utility of alternative predictors of job performance*." Psychological bulletin, *96*(1), 72-98, (1984).
[14] M. A. McDaniel, F. L. Schmidt and J. E. Hunter. "*Job experience correlates of job performance*." Journal of Applied Psychology, 73, 327-330, (1988).
[15] M. A. Quinones, J. K. Ford, J. and M. S. Teachout. "*The relationship between work experience and job performance: A conceptual and meta-analytic review*," Personnel Psychology, *48*, 887-910, (1995).
[16] R. H. Adlam. "The US oil and gas industry – a safety culture enigma. Offshore Network Limited, England and Wales (2013).
[17] A. Williamson, D. A. Lombardi, S. Folkard, J. Stutts, T. K Courtney, and J. L. Connor. "*The link between fatigue and safety*." Accident Analysis & Prevention, *43*, 498-515. (2011).
[18] S. Folkard and D. A. Lombardi. "*Modeling the impact of the components of long work hours on injuries and "accidents"*, American Journal of Industrial Medicine, *49*, 953-963, (2006).
[19] F. W. Guldenmund. "*(Mis) understanding Safety Culture and Its Relationship to Safety Management*." Risk analysis, *30*, 1466-1480, (2010).
[20] B. Graham, W. K. Reilly, F. Beinecke, D. F. Boesch, T. D. Garcia, C. A. Murray, and F. Ulmer. "*The National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling*." Deep Water. The Gulf Oil Disaster and the Future of Offshore Drilling. *Report to the President*, (2011).
[21] E. Salas, D. E. Sims and C.S. Burke. "*Is there a "Big Five" in teamwork?"* Small group research, 36, 555-599, (2005).
[22] C. K. De Dreu and L. R. Weingart. "*Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis*." Journal of applied Psychology, 88, 741-749 (2003).