Safety Science 78 (2015) 117–123

Contents lists available at ScienceDirect

Safety Science

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

Psychophysiology, task complexity, and team factors determine emergency response teams' shared beliefs

Bjørn Sætrevik *

University of Bergen, Norway

ARTICLE INFO

Article history:
Received 16 September 2014
Received in revised form 15 March 2015
Accepted 14 April 2015
Available online 2 May 2015

Keywords:
Situation awareness
Shared mental models
Team cognition
Field experiment
Emergency response

ABSTRACT

In field settings where the objective truth is not known, the extent to which you have the same understanding of the situation as your team leader may be used as an indicator for a team's situation awareness. Two experiments showed emergency response team members' degree of shared beliefs (measured as a 'similarity index') to be associated with which team they are in, but not with which position they have in the team. This indicates that factors specific to the teams, e.g. the leader's behavior, the team's shared experience, or communication patterns, are important for a team's situation awareness. In the second experiment, task complexity was manipulated with a scripted scenario design and heart rate variability was measured as an indicator of executive function. Shared beliefs were shown to be associated with the degree of high frequency modulation of heart rate variability. Further, shared beliefs were associated with the designed task complexity for some teams. The experiments showed no association between the measure of shared beliefs and subjective reports of situation awareness.

1. Introduction

1.1. Mental representations in safety team-work

Safety critical work and work in high reliability organizations is often done in teams, where two or more operators with different responsibilities and skill sets cooperate toward shared goals (Salas et al., 1992). The organization of operators into teams may be due to workload demands or due to requirements for diverse skill sets. There may be advantages to performing safety critical work in teams compared to doing it as individual work, as the different team members can monitor and assist each other to achieve higher reliability. However, organizing the work in teams also carries disadvantages, such as creating a more complex work environment, losing resources to imperfect communication and coordination, and the risk for uncoordinated team members working toward opposing goals.

In order for team organization and team training to reduce adverse team effects, there is a need for research on the factors that influence a team's efficiency and safe functioning. Relationships suggested by a-priory theorizing or correlational findings (e.g. Eid et al., 2011; Gross and Kluge, 2012; Guldenmund, 2000; Kanno et al., 2013) should be tested in experimental designs. In order to ensure the applied value of the research, studies should strive for both ecological validity and controlled hypothesis testing.

A fundamental assumption when studying team reliability is that individual and aggregated task performance improves when the team members have accurate mental representations of the situation. Further, a team where all the members have accurate representations will also have similar (or shared) representations. This should facilitate communication and cooperation, and thus increase the team's overall performance (Cannon-Bowers and Salas, 2001; Mathieu et al., 2000; Sætrevik and Eid, 2014; Salas et al., 2008; Saner et al., 2009).

A number of theoretical concepts and measurement approaches have been suggested to describe the accuracy and cohesion of team's mental representations. *Mental models* are organized and dynamic internal representations of past experiences (Glaser, 1989), and individual team members will have mental models describing their tasks and the team's work. The mental models could to some extent be similar between individual members of a team, indicating that they have the same knowledge or assumptions about the situation. Salas and colleagues (Cannon-Bowers et al., 1993; Salas et al., 1992) have referred to this phenomenon as a team having *shared mental models* (SMM). The content of the SMMs may be task-related (e.g. relating to the equipment or job strategies) or team-related (information about e.g. interaction patterns in the team, or the skills of various team members, Mathieu

* Address: Faculty of Psychology, University of Bergen, Christies gate 12, NO-5015 Bergen, Norway.

E-mail address: satrevik@gmail.com

http://dx.doi.org/10.1016/j.ssci.2015.04.017
0925-7535/© 2015 The Author. Published by Elsevier Ltd.

et al., 2000). Salas et al. (2007) argued that a higher degree of SMM is indicated by a team using closed-loop communication, performance monitoring between team members, and displaying supportive behavior.

The term *situation awareness* (SA) refers to the extent to which an individual's mental representation of the dynamic environment corresponds to the actual environment. SA also refers to the process that creates this representation in an interaction with the environment. The prevailing model for SA divides the concept into three hierarchical levels, which consist of perceiving, understanding, and predicting the environment (Endsley, 1988a, 1995). To account for SA in teams, the concept *team SA* has been introduced to describe the aggregation of individual team members' accurate mental representations of their own fields of responsibility, while *shared SA* has been introduced to describe the degree to which all team members have accurate mental representations of issues that are relevant for the whole team (Endsley, 1995). This classification emphasizes that some information in a team's work needs to be shared, while other information does not.

1.2. Relevance of mental representation accuracy for emergency response team-work

To organize work into teams may have effects that are adverse for safety, for example that the team members misunderstand instructions, work toward opposing goals, fail to utilize all the team's information or resources, get involved in interpersonal conflicts or social loafing. For a high reliability organization such as the emergency response teams (ERT) of offshore hydrocarbon energy industry, well-coordinated team-work is critical to mission success. The members of such teams may have other tasks and work teams in their day-to-day jobs, but have to muster to the ERT in case of emergencies. In some cases the emergency preparedness approach may be organized into a first-line (or frontline) operational ERT, a second-line tactical ERT, and a third-line strategic ERT. While the first-line ERT directly interacts with the event causing the emergency, the second-line ERT is tasked with collecting and organizing all information relevant for the event, relaying information between parties involved in the event, planning for future development of the event, and advising the first-line ERT.

Everyday safety (e.g. avoiding mistakes, errors and slips during normal operation, Reason, 1990) may be determined by other factors than those important for maintaining safety while dealing with an emergency. The task-work involved for an ERT trying to normalize an emergency is done while team members are aware of the high stakes involved, their physiological activation may be increased, the teams may have limited experience in their emergency tasks and with working together, and the task to be solved may be novel. Thus an ERT's work may be especially prone to erroneous actions due to inaccurate mental representations for individual team members or due to the team members having non-shared mental models. Further, the team-work errors may be more difficult to notice and the consequences of team-work errors may be larger than during normal operations.

1.3. Determinants for mental representation accuracy

One may propose three domains of determinants for the extent to which team members have accurate representations about their environment: the individual, the team and the task. Individual team members vary in their level of competency and skill (see Gross and Kluge, 2014, for an applied example), which may allow them different baselines for learning about and understanding their environment. Further, one may expect individual variation in affective and motivational aspects, which may be reflected in psychophysiological activation and the capacity for mental representation (Gonzalez, 2005; Gonzalez et al., 2005). To some extent, the organization may enhance such individual factors through personnel selection and training, and through influencing the cultural safety values. In a research design where team or task factors are examined (as in the one presented here), the contribution of individual factors would create noise in the analysis and would be averaged away. On the other hand, if the mental representations are predominantly determined by individual factors, the analyses would not find significant contributions when analyzing team or task factors.

In addition to team members developing mental representations through their individual information gathering and hypothesis testing, there is also sharing, discussion and organization of information between team members. This makes the team-level a relevant domain for determinants for mental representation accuracy. The concept of *team cognition* (Fiore and Salas, 2004; Salas et al., 2007) has been used to describe multi-level interactions and dependencies between intraindividual and interindividual processes. Team cognition is seen as analogous to individual cognition, and is an emergent state where important information is organized, represented and distributed, which allows anticipation and execution of the task (Kozlowski and Ilgen, 2006). Different traditions tend to measure team cognition either as compositional representation (i.e. shared SA or SMM) or as compilational representations (transactive memory systems). Salas et al. (2007) argued that SMM could be a reliable marker for team cognition, which would allow the team members to make accurate causal explanations and adapt efficiently to each other. Meta-analyses have indicated that team cognition has strong positive relationships to behavior, motivation and performance (DeChurch and Mesmer-Magnus, 2010; Mathieu et al., 2000, 2010). One may expect different teams to differ in their capacity for team cognition, which could cause differences in the accuracy of team members' representations and hence of the team's aggregated representations. It has been argued (DeChurch and Mesmer-Magnus, 2010) that factors such as team leadership, shared experience and training, and workplace design may enhance team cognition. A related concept is *macrocognition in teams*, which tends to emphasize collective knowledge building in novel situations (Fiore et al., 2010). This research tradition originated in naturalistic decision-making research, and parallels distributed cognition literature. The concepts of team cognition and macrocognition both correspond to the idea that an effective team works with a high degree of coordination due to team members having the same understanding of the situation and task goals.

A third domain of predictors for a team's accurate mental representations is aspects of the task the team is working on and how the team members are situated in the task (Gonzalez, 2005). A task may vary in its opaqueness, task complexity, and dynamic complexity (Diehl and Sterman, 1995; Hardman, 2009). As a task grows more complex it allows for more degrees of freedom in how the mental representations are structured (Wood, 1986), which increases the potential for inaccurate mental representation. Further, given that organizing work into teams always involves a specialization of tasks and competencies (as per the definition by Salas et al. (1992)), a team member holding a given position in the team will have a different access to and a different perspective on information about the task and the team's work than a team member in a different position. As revealed in the discussion of *team SA* and *shared SA* in Section 1.1, some information should be shared by team members in all positions, while some information can be exclusive to only some positions. If team members have different mental representations of the information that all positions should share, this can lead to reduced team performance and risk for human error. As for the information that is not shared in the team, a team member is likely to have more accurate representations of information that is more pertinent to her or his team role's task domain. Additionally, a more generalized or supervisory role in the team may give a superficial access to many task domains, whereas a more specialized role gives detailed access to a single task domain.

1.4. Measurement of mental representation accuracy

Concepts such as SA and SMM are clearly important for understanding safe and efficient teamwork (see e.g. Woods and Sarter, 2010), but the correspondence between inner states and external reality has proven difficult to measure, and different approaches to measuring the accuracy of mental representations have been suggested. Some measures are *subjective* (e.g. Taylor, 1990), in that they ask the operators to evaluate the accuracy of their mental representations. Other measures are *objective* (e.g. Endsley, 1988b), in that they rely on comparing what the operators believe to be true about the situation with the experimenter's knowledge about the reality. In some cases *process measures* (see e.g. Patrick and Morgan, 2010) may be used to assess the operator's cognitive processes based on her or his actions, e.g. by examining communication, eye-movements or interaction with a control system. Finally, one may use *observers* to rate the accuracy of operator's beliefs based on their observable actions (e.g. Matthews et al., 2000). See other publications (e.g. Hone et al., 2006; Sætrevik and Eid, 2014; Salmon et al., 2006; Saner et al., 2009) for more detailed reviews of different measurement techniques and their limitations. The preferred approach to measure mental representation accuracy may partly depend on the research setting affordances, and partly on the researcher's theoretical understanding of SA. Salmon et al. (2006) recommended that studies of SA should combine a variety of measures.

In applied settings where no standard for objective truth is available or can feasibly be measured, it may be challenging to measure the cohesion of a team's knowledge states (i.e. shared SA or SMM). Thus, while mental representation accuracy is likely to predict safe team performance, a precise measurement of it is impossible in all but the most artificial settings. I have previously argued (Sætrevik and Eid, 2014) that in applied settings it may be meaningful to measure the individual team members' beliefs about the situation, and use the extent to which they share the same beliefs to calculate a *similarity index*. This builds on the assumption that although individual beliefs will depart from the objective truth, an "average" of the measurements may be closer to the truth. An approach of this type has also been suggested by Saner et al. (2009).

If the setting allows one to assume that one of the team members, e.g. the team leader, should on average be better informed than the rest of the team, one may compare the answers from the rest of the team to the answers of this member, and assume that a higher degree of agreement represents more accurate knowledge. If the team leader is indeed more accurate than other team members, the similarity between a team member's beliefs and the team leader's beliefs may approach a measure of SA, despite not directly measuring whether the beliefs correspond to an objective reality. Thus the similarity of a team member's beliefs to her or his team leader's beliefs will be used as a proxy for having accurate beliefs about the work and the team, which should correspond to high levels of SA. I will refer to this approach as measuring the degree of *shared beliefs* of a team or of a team member. In the terms of DeChurch and Mesmer-Magnus (2010), this is a measure of compositional emergence. A more detailed description of the similarity index approach, inherent assumptions and caveats can be found in a previous publication (Sætrevik and Eid, 2014).

To function in a complex environment, an organism must balance between and modulate control from the sympathetic nervous system, involved in *flight-or-fight* responses, and the parasympathetic nervous system, involved in rest and restitution (Thayer et al., 2012). It is assumed that an alert and well-adapted organism continually adjusts its bodily activation to suit the current environment demands, whereas an organism that is under high levels of emotional stress to less degree does these moment-to-moment calibrations and its behavior is more regulated by automatic and prepotent processes. The degree of high frequency modulation of heart rate may reflect the organism's modulation of cognitive systems to adapt to the environment. It has been argued (Luft et al., 2009; Thayer et al., 2009) that the central autonomic network links the activation in prefrontal cortex neural function to heart rate variation (HRV). Thus the degree of high-frequency modulation of HRV can be used as an indicator of executive function. It has been showed that indices of HRV positively correlates with performance on working memory and executive function tasks (Elliot et al., 2011; Hansen et al., 2003) and with field measures of SA (McKneely et al., 2006; Saus et al., 2012).

1.5. Current study

In the current research project, I wanted to measure the similarity index of ERT members in an ecological setting in order to identify some of the factors that lead to shared mental representations which improves the organization's safety. A further research interest was to assess the validity of subjective estimates of SA in the current setting. A final research interest was to assess whether the measurements reflect changes in cognitive states induced by changes in the team's task. To move beyond self-report and similarity measures of mental representation and toward a more direct measure of internal states, HRV was used as a psychophysiological measure of executive function (Hansen et al., 2003). This allows for testing of whether the task manipulation, similarity index and subjective reports reflect changes in internal states.

We performed two field experiments among the six ERTs of a hydrocarbon energy company, as described in Section 1.1. For more details on the setting for the data collection, please see a previous publication (Sætrevik and Eid, 2014). Each of the six teams had around ten members. Both experiments had the same measures of shared beliefs and of subjective SA. Experiment 2 had a higher degree of experimental control, more frequent measurements, manipulated the task complexity, and recorded psychophysiological activation (HRV).

1.6. Hypotheses

Based on the possible determinants of mental representations in teams discussed in Sections 1.3 and 1.4, the current study will test the following hypotheses: Given that team factors may cause members of some teams to have more accurate mental representations than members of other teams, the similarity index will vary between teams (H1). Further, if the team builds their understanding of the situation over the time they collaborate on a task, the similarity index will increase over scenario time (H2). Given that a complex task offers more degrees of freedom for how a mental representation can be structured, shared beliefs will decrease with task complexity, yielding lower similarity index (H3). Recent research has shown task induced psychophysiological activation to be associated with decreased executive function. Thus shared beliefs (as indexed by similarity index) will decrease when cognitive resources (as indexed by HRV) decreases (H4). Finally, given that SA is often measured with subjective assessment, one may expect the subjective SA measure to predict the level of similarity index (H5). Descriptive statistics for both experiments are provided elsewhere (Sætrevik and Eid, 2014).

2. Experiment 1

2.1. Experiment 1 methods

2.1.1. Experiment 1 sample

The participants in both experiments were second-line ERTs that are tasked with gathering and organizing information about a dynamically evolving external event, such as a fire on an offshore oil rig. Team members largely work individually, collecting information from external human, document and instrument sources, and their work is organized by a team leader. All the six ERTs in the organization participated in the data collection. There were one team leader and between 9 and 11 team members on each team, yielding a total of 58 participants.

2.1.2. Experiment 1 procedure

Data collection for Experiment 1 was done during the ERT's regular training exercises. The training exercises are run as realistically as possible, with the ERT located in the emergency rooms, using the actual equipment, and interacting with the onshore and offshore personnel that would be involved in the handling of an actual incident of this type at this time. Experiment 1 placed no restrictions on the scenario design, which lasted between 2 and 3 h.

During scenario play, at the time when the team leader announced a status meeting, team members were told to answer eight probe questions in a pen-and-paper booklet. The first five probe questions queried for information about the task and the team's work, asking about the incident location, incident type, status of personnel, and asked team members to predict the outcome of the incident and to state the current priorities in the team's work. The similarity index was computed for each probe question by calculating the numerical distance between the team member's answer and the team leader's answer, and standardizing it to vary from 0 to 1. Similarity index scores approaching 1 on probe questions indicated that a team member had given the same answer as the team leader, while lower similarity scores indicated more difference between the team member's and the team leader's answers. For more details on the method of calculation, please see a previous publication (Sætrevik and Eid, 2014). The final three probe questions were inspired by the 3D-SART measure (Taylor, 1990), and asked about the team member's subjective evaluation of their SA in terms of their access to the information they needed, the demand on their attention and the completeness of their understanding for their area of responsibility.

2.1.3. Experiment 1 analyses

The analyses will use team members' average scores of similarity index (across probes 1–5) to measure shared belief, and average subjective SA scores (across probes 6–8). For a discussion of score distribution across probes 1–5, see a previous publication (Sætrevik and Eid, 2014). The ERTs had between 4 and 6 status meetings during the scenario. In order to compare the teams, only the first four data points from each team were used in the analysis.

Two repeated measures general linear model (GLM) analyses, using the least square methods, were performed to test the effects on shared beliefs and on subjective SA separately. The first GLM had team membership and team position as categorical predictors, subjective SA score as continuous predictor, and four repeated measures of shared beliefs as outcome variable. If a participant missed less than half of the shared belief measurements, the missing values were replaced with the participant's average of the remaining scores (24 of 184 data points imputed). A second GLM had team membership and team position as categorical predictors, scenario average shared beliefs as continuous predictor, and four repeated measures of subjective SA as outcome variable. If a participant missed less than half of the four subjective SA measurements, the missing scores were replaced with the participant's average of the remaining scores (30 of 208 data points were imputed). Applying Bonferroni correction for three predictors yields an alpha level of $p = .0167$ (0.05/3) for both GLMs. All analyses were done using statistical package STATISTICA 12 (StatSoft Inc., Tulsa, OK, USA).

2.2. Experiment 1 results and discussion

The first GLM showed a significant effect of team membership ($F(5,38) = 3.67$, $p = .008$) showing that shared belief was predicted by team membership, in the sense that the members of some teams scored higher for shared belief than the members of other teams. There was no effect of team position ($F(7,38) = .81$, $p = .59$), of subjective SA ($F(1,38) = 1.27$, $p = .27$), or of measurement time ($F(3,114) = .65$, $p = .58$). None of the interactions were significant. The interaction closest to significance was between team membership and time of measurement ($F(15,114) = 1.56$, $p = .1$), indicating that different teams may have had different trajectories across the scenario, but there was no clear trend of teams increasing or decreasing their shared belief over time.

The second GLM showed no significant main effects of team membership $F(5,38) = .34$, $p = .88$), of role ($F(7,38) = 1.99$, $p = .08$), of shared beliefs ($F(1,38) = 0.46$, $p = .5$) or of measurement time ($F(3,114) = 1.19$, $p = .32$). Thus the participants' subjective rating of their SA was not predicted by which team they belonged to, which position they had in the team, or the degree to which they had the same beliefs as their team leader. There was a non-significant trend of increased SA confidence across the four measurement points (mean = 5.07, $SD = 0.17$, mean = 5.55, $SD = 0.12$, mean = 5.7, $SD = 0.12$, mean = 5.73, $SD = 0.12$). There was a significant interaction effect between time of measurement and team membership ($F(15,114) = 2.43$, $p = .004$), indicating that different teams rated their SA differently at different time points. When examining the means, three of the teams showed a trend of increasing their estimates as the scenario progressed while the other three teams did not have a clear trend. The interaction between time of measurement and team position ($F(21,114) = 1.75$, $p = .04$), which would indicate that SA confidence changed over time for some team positions, did not satisfy the Bonferroni corrected alpha level.

3. Experiment 2

3.1. Experiment 2 methods

3.1.1. Experiment 2 sample

The data collection of Experiment 2 aimed for a higher degree of experimental control, fewer missing values, psychophysiological measurement and experimental manipulation of task complexity in order to draw causal inferences. The same participants took part in Experiment 2 as in Experiment 1, but organized into different teams. As not all of the ERT personnel in the organization were able to participate this time, only five teams took part, with one team leader and eight team members in each team (for a total of 45 participants in Experiment 2).

3.1.2. Experiment 2 procedure

A detailed scenario was scripted, and was executed using a staff of actors playing the roles of the ERT's counterparts in all scenario communications. Further, in order to get better temporal resolution and to disentangle the measured variables from the team leader's decision to arrange status meetings, eight measurement points were planned at set times, at 20, 40, 60, 80, 100, 120, 140, and 160 min after scenario start (participants were not informed about the measurement schedule). To facilitate data-processing, electronic questionnaires on the team member's work station were used, while the probe content was mostly similar to Experiment 1 (for details, see Sætrevik and Eid, 2014). In order to separate the effect of time passed from other factors that could have an effect on the variation in task performance, the scripted scenario was designed to have an ABAB variation of complexity. In this design, the task was more complex in the first and third quarter of the scenario run, and more straight-forward in the second and fourth quarter of the scenario run. Complexity was manipulated according to the descriptions of Campbell (1988) and Wood (1986).

In order to have an accessible indicator of psychophysiological activation, all participants in Experiment 2 were outfitted with consumer grade heart rate measurement equipment (Polar RSX800CX, consisting of a chest sensor band and a wrist receiver unit). The time of onset for each heartbeat (R-wave to R-wave, RR) was measured continuously throughout the scenario. To avoid overestimating the impact of sympathetic influence due to non-stationarity in the recording (Jorna, 1992; Magagnin et al., 2011), eight segments containing the 5 min preceding each questionnaire measurement were extracted for analysis. From these values, the root mean sum of squared distances (rMSSD; Berntson et al., 2005) of RR intervals in heart rate were calculated to get a single value indicator that relates to high-frequency in modulating the heart rate.

3.1.3. Experiment 2 analyses

Three repeated measures GLMs were performed to test the effects on shared beliefs, on subjective SA, and on HRV separately. The first GLM had team membership, team position and task complexity as categorical predictors, scenario average subjective SA score and scenario average HRV as continuous predictors, and eight repeated measures of shared beliefs as outcome variable. If a participant missed less than half of the eight shared belief scores the missing scores were replaced with the participant's average of the remaining scores (25 of 320 data points were imputed). The second GLM had team membership, team position and task complexity as categorical predictors, scenario average shared beliefs and scenario average HRV as continuous predictors, and eight repeated measures of subjective SA as outcome variable. If a participant missed less than half of the eight subjective SA scores, the missing values were replaced with the average of the participant's remaining scores (36 of 360 data points were imputed). The third GLM had team membership, team position and task complexity as categorical predictors, scenario average shared beliefs and subjective SA as continuous predictors, and eight repeated measures of HRV as outcome variable. If a participant missed less than half of the HRV measurements (due to signal loss), the missing values were replaced with the average of the participant's remaining measurements (4 of 320 data points were imputed). Applying Bonferroni correction for five predictors yields an alpha level of $p = .01$ (0.05/5) for all three GLMs.

3.2. Experiment 2 results and discussion

The first GLM showed a main effect of team membership ($F(4,19) = 13.33$, $p < .001$), indicating that members of some teams had higher degree of shared beliefs than members of other teams. There was also a main effect of HRV ($F(1,19) = 9.4$, $p < .001$), indicating that having more executive function resources available (as indexed by rMSSD of HRV) was associated with higher degree of shared beliefs. There was no main effect of team position ($F(4,19) = 2.14$, $p = .09$) or of task complexity ($F(1,19) = 2.64$, $p = .12$). There was a significant interaction effect between team membership and task complexity ($F(4,19) = 6.29$, $p = .002$), indicating that some teams had less shared belief in the parts of the scenario with higher complexity. There was also a significant interaction between team membership, task complexity and measurement time ($F(12,57) = 3.72$, $p < .001$), indicating that different teams had different shared belief trajectories over time relative to task complexity conditions.

The second GLM had no significant main effects. There was a significant interaction between team membership and task complexity ($F(4,19) = 5.5$, $p = .004$), indicating that some teams expressed more confidence in the lower complexity condition, while others did not. The interaction between team membership and measurement time ($F(12,57) = 2.26$, $p = .02$), indicating that some teams increased their confidence over time, while others did not, did not satisfy the Bonferroni corrected alpha level. Inspecting the means showed that four teams increased their SA confidence through the experiment, while the final team had lower confidence on the last measurements than on the first. There was a significant interaction between team membership, task complexity and measurement time ($F(12,57) = 3.02$, $p < .003$), indicating that some teams increased their confidence across scenario time relative to the task complexity, while other teams did not.

The third GLM showed a main effect of average shared belief ($F(1,19) = 5.69$, $p = .003$), indicating that the team members who showed more psychophysiological adaptability to the environment (indicating executive function resources) were also the participants with higher degree of shared beliefs. There were no main or interaction effects of team membership, team position, subjective SA or time of measurement.

4. General discussion

4.1. Summary of results

In the present two experiments, team members' shared beliefs (i.e. their similarity index score) varied more between teams than within teams. This indicates that some aspect of the team membership accounted for the degree to which team members had the same understanding of the situation as their team leader. There was also an interaction in Experiment 2 indicating that task complexity accounted for the variation in shared beliefs for some of the teams. However, shared beliefs did not vary consistently over time, indicating that the participants did not consistently come closer to their team leader's understanding of the situation as the scenario progressed. Despite this, there was a trend for increasing subjective evaluation of SA, indicating that some teams became more confident over time. Importantly, in neither experiment was subjective evaluation of SA associated with the degree of shared beliefs.

Further, Experiment 2 measured psychophysiological activation in terms of HRV. A greater degree of modulation of activation was taken as an indicator for the team member's executive function in adapting to a changing task environment. Results showed HRV to be associated with scenario average similarity index, indicating that team members with higher degrees of shared beliefs showed more adaptability to the environment. Conversely, the team members' average subjective SA did not account for the variation in HRV. The HRV did not vary consistently over the time course of the scenario, and did not vary according to the designed task complexity.

Some of the hypotheses presented in Section 1.6 were supported while others were not. H1 was supported, as team membership predicted shared beliefs. H2, that shared beliefs would increase over scenario time was not clearly supported, although there was a trend for an interaction with team membership in Experiment 1, and an interaction with team membership and complexity in Experiment 2. H3 was partly supported, as Experiment 2 showed that shared beliefs decreased with task complexity for some teams. H4 was supported, as shared beliefs decreased when psychophysiological indicator for executive function (HRV) decreased. Finally, there was no support for H5, as subjective assessment of SA was not associated with the degree of shared beliefs.

4.2. Mechanisms for developing shared beliefs

Both experiments indicated that team membership predicts more accurate mental representations (in terms of sharing the beliefs of your team leader). However, the experiments are not suited to inform us about what it is about some teams that gave them more accurate mental representations. As discussed in Section 1.2, it is appropriate to assume that the patterns in compilational mental representations may be due to differences in team cognition or macrocognition between the teams. It has been argued (DeChurch and Mesmer-Magnus, 2010; Fiore et al., 2010; Salas et al., 2008, 2007) that in order to be useful, such concepts should be tied to behavior markers and to the mechanisms that causes behavior to lead to enhanced team coordination and team effectiveness.

One possible mechanism is that acts of leadership has a direct effect on the accuracy of the mental representation of the team members. In the current setting, the team leader is tasked with organizing the team's work, collecting and sharing information and setting team priorities. The team leader does this mostly through one-on-one communication directly to individual team members, and through brief plenary status meetings arranged at the leader's discretion about every half hour. Previous analyses of this data set (Sætrevik and Eid, 2014) showed that the frequency of status meetings or the time passed since previous meeting did not predict the degree of shared beliefs. Thus, if the leader's communication is critical, it must be more subtle variations in the communication that determines team performance. In further studies, it could be examined whether the team differences in shared beliefs are due to the quality or the content of the leader's communication by analyzing the content of one-on-one conversations or status meetings. It could also be of interest to examine the pattern of communication and relationships within the teams, to see if teams with high degrees of shared beliefs have different communication structures or social structures than teams that do not function as well. In order to answer questions on this level of analysis, a more detailed data collection will be needed where minute actions and information flow are studied. An example of this would be the *TeamPrints* approach (Bolstad et al., 2007).

It could be the team leader's overall leadership style that influences the team's performance, for example by enhancing the team members' motivation or the intra-team relationships (i.e. *authentic leadership*, Avolio and Gardner, 2005, or *transformational leadership*; Eid et al., 2004). Alternatively, the team leader's actions may directly influence the team's information flow, through monitoring and directing team members' work and attention to where it is needed. Such hypotheses could be tested by examining the effect on shared beliefs of replacing the team's leader, using a similar design as used by Espevik et al. (2006).

4.3. Implications

The current study found mental representations' accuracy (measured as similarity to team leader's beliefs) to be stable across different team roles, but to vary between teams. This indicates that factors specific to the teams determine the team members' ability to orient themselves in the situation. This appears to justify further work on identifying such factors and on improving them in training programs. Manipulating the task scenario showed that for some teams the mental representations' accuracy suffered when task complexity increased. It could be that team members can be trained to counter the effects of information complexity, for example by learning to structure the way information is communicated, and case-based training programs could help team members to quickly identify typical situations and structure incoming information around appropriate cognitive schemas.

The current study successfully applied a novel field measure (see Sætrevik and Eid, 2014, for details) for testing hypotheses about a team's mental representation accuracy, which indicates that the approach may also be applied in future studies. Further, the similarity index as a measure of shared beliefs in a team was a more useful approach to measure the accuracy of a team member's mental representations than self-reported SA was. This may support Endsley's claim (Endsley, 1994) that the utility of subjective SA measurements is limited, as they are an indicator of the responder's confidence, rather than indicating the accuracy of their beliefs about the situation. The study thus indicates that it is not advisable to use subjective evaluation of SA as an indicator for the accuracy of mental representations. The measurement approach may be of interest for researchers examining the factors contributing to the development and maintenance of SMM or SA. The approach may also be implemented by industry trainers that seek to identify team members, team positions or task aspects where SA declines and additional resources are needed. When studying several teams in the same setting, the approach can be used to identify differences between teams (e.g. different communication patterns) that cause variation in SA.

The current study found that team members' physiological activation was associated with their degree of shared beliefs. To my knowledge, HRV has not previously been associated with SA for ERTs. As heart rate measurement is relatively manageable in field settings, it may be a fruitful as an objective measure of executive function among team members in future field studies of team performance and for training. Task complexity was seen to interact with team membership in modulating HRV, and to be associated with the degree of shared beliefs. It thus appears appropriate that teams should be aware of the impact of psychophysiological activation on team performance.

Acknowledgements

I appreciate the assistance of Evelyn-Rose Saus, Sigurd William Hystad, Gry Dammen Tyssebotn, Annette Marie Kollenborg, Astrid Lovise Westvik and Daniel Hosøy in the data collection. Leif Arvid Øvernes assisted in the heart rate variability analysis, while Anita Lill Hansen was a helpful discussion partner on this issue. The second line emergency response organization of Statoil's Development and Production Norway, Logistics and Emergency preparedness served as our industry partner and setting for data collection, and contributed in designing the probe items and running the scenarios. This research was supported by the Norwegian Research Council's PETROMAKS program (Grant Number 189521). An article describing the methodology in more detail (Sætrevik and Eid, 2014) has been published, but the hypotheses presented here have not previously been tested on these datasets.

References

Avolio, B.J., Gardner, W.L., 2005. Authentic leadership development: getting to the root of positive forms of leadership. Leadership Quart. 16, 315–338.

Berntson, G.G., Lozano, D.L., Chen, Y.J., 2005. Filter properties of root mean square successive difference (RMSSD) for heart rate. Psychophysiology 42, 246–252.

Bolstad, C.A., Foltz, P., Franzke, M., Cuevas, H.M., Rosenstein, M., Costello, A.M., 2007. Predicting situation awareness from team communications. Hum. Factors Ergon. Soc., 789–793.

Campbell, D.J., 1988. Task complexity: a review and analysis. Acad. Manage. Rev., 40–52.

Cannon-Bowers, J.A., Salas, E., 2001. Reflections on shared cognition. J. Org. Behav. 22, 195–202.

Cannon-Bowers, J.A., Salas, E., Converse, S., 1993. Shared mental models. In: Castellan, J.N. (Ed.), Individual and Group Decision Making. Erlbaum, Hillsdale, NJ, pp. 221–246.

DeChurch, L.A., Mesmer-Magnus, J.R., 2010. The cognitive underpinnings of effective teamwork: a meta-analysis. J. Appl. Psychol. 95, 32.

Diehl, E., Sterman, J.D., 1995. Effects of feedback complexity on dynamic decision making. Organ. Behav. Hum. Decis. Process. 62, 198–215.

Eid, J., Johnsen, B.H., Brun, W., Laberg, J.C., Nyhus, J.K., Larsson, G., 2004. Situation awareness and transformational leadership in senior military leaders: an exploratory study. Military Psychol. 16, 203.

Eid, J., Mearns, K., Larsson, G., Laberg, J.C., Johnsen, B.H., 2011. Positive organizational behaviour and safety science: conceptual issues and future research questions. Saf. Sci. 50, 55–61.

Elliot, A.J., Payen, V., Brisswalter, J., Cury, F., Thayer, J.F., 2011. A subtle threat cue, heart rate variability, and cognitive performance. Psychophysiology 48, 1340–1345.

Endsley, M.R., 1988a. Design and evaluation for situation awareness enhancement. In: Proceedings of the Human Factors Society 32nd Annual Meeting. Human Factors Society, Santa Monica, CA, pp. 97–101.

Endsley, M.R., 1988b. Situation awareness global assessment technique (SAGAT). Proc. Acad. Nat. Sci. Phila., 789–795.

Endsley, M.R., 1994. Situation awareness in dynamic human decision making: measurement. In: Gilson, R.D., Garland, D.J., Koonce, J.M. (Eds.), Situational Awareness in Complex Systems. Embry-Riddle Aeronautical University Press, Daytona Beach, FL, pp. 79–97.

Endsley, M.R., 1995. Toward a theory of situation awareness in dynamic systems. Hum. Factors: J. Hum. Factors Ergon. Soc. 37, 32–64.

Espevik, R., Johnsen, B.H., Eid, J., Thayer, J.F., 2006. Shared mental models and operational effectiveness: effects on performance and team processes in submarine attack teams. Military Psychol. 18, S23–S36.

Fiore, S.M., Salas, E., 2004. Why we need team cognition. Team Cognition: Understanding the Factors that Drive Process and Performance. American Psychological Association, Washington, DC, pp. 235–248.

Fiore, S.M., Rosen, M.A., Smith-Jentsch, K.A., Salas, E., Letsky, M., Warner, N., 2010. Toward an understanding of macrocognition in teams: predicting processes in complex collaborative contexts. Hum. Factors: J. Hum. Factors Ergon. Soc. 52, 203.

Glaser, R., 1989. Expertise and learning: how do we think about instructional processes now that we have discovered knowledge structures. In: Klahr, D., Kotovsky, K. (Eds.), Complex Information Processing: The Impact of Herbert A. Simon. Erlbaum, Hillsdale, NJ, pp. 269–282.

Gonzalez, C., 2005. Task workload and cognitive abilities in dynamic decision making. Hum. Factors: J. Hum. Factors Ergon. Soc. 47, 92–101.

Gonzalez, C., Thomas, R.P., Vanyukov, P., 2005. The relationships between cognitive ability and dynamic decision making. Intelligence 33, 169–186.

Gross, N., Kluge, A., 2012. ''Why should I share what I know?''-antecedents for enhancing knowledge-sharing behavior and its impact on shared mental models in steel production. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, pp. 403–407.

Gross, N., Kluge, A., 2014. Predictors of knowledge-sharing behavior for teams in extreme environments an example from the steel industry. J. Cognit. Eng. Decis. Making, 1555343414540656.

Guldenmund, F.W., 2000. The nature of safety culture: a review of theory and research. Saf. Sci. 34, 215–257.

Hansen, A.L., Johnsen, B.H., Thayer, J.F., 2003. Vagal influence on working memory and attention. Int. J. Psychophysiol. 48, 263–274.

Hardman, D., 2009. Judgment and Decision Making: Psychological Perspectives. John Wiley & Sons, Chichester, UK.

Hone, G., Martin, L., Ayres, R., 2006. Awareness – does the acronym ''SA'' still have a practical value? In: 11th International Command and Control Research and Technology Symposium, Cambridge, UK.

Jorna, P., 1992. Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. Biol. Psychol. 34, 237–257.

Kanno, T., Furuta, K., Kitahara, Y., 2013. A model of team cognition based on mutual beliefs. Theor. Iss. Ergon. Sci. 14, 38–52.

Kozlowski, S.W., Ilgen, D.R., 2006. Enhancing the effectiveness of work groups and teams. Psychol. Sci. Public Interest 7, 77–124.

Luft, C.D.B., Takase, E., Darby, D., 2009. Heart rate variability and cognitive function: effects of physical effort. Biol. Psychol. 82, 186–191.

Magagnin, V., Bassani, T., Bari, V., Turiel, M., Maestri, R., Pinna, G.D., Porta, A., 2011. Non-stationarities significantly distort short-term spectral, symbolic and entropy heart rate variability indices. Physiol. Meas. 32, 1775.

Mathieu, J.E., Heffner, T.S., Goodwin, G.F., Salas, E., Cannon-Bowers, J.A., 2000. The influence of shared mental models on team process and performance. J. Appl. Psychol. 85, 273–283.

Mathieu, J.E., Rapp, T.L., Maynard, M.T., Mangos, P.M., 2010. Interactive effects of team and task shared mental models as related to air traffic controllers' collective efficacy and effectiveness. Hum. Perform. 23, 22–40.

Matthews, M.D., Pleban, R.J., Endsley, M.R., Strater, L.D., 2000. Measures of infantry situation awareness for a virtual MOUT environment. In: Proceedings of the Human Performance, Situation Awareness and Automation: User Centred Design for the New Millennium Conference.

McKneely, J., Bevan, M., Cropper, K., Iny, M., Vaughan, F., 2006. Initial Investigation on Fatigue in Command and Control Situation Awareness: Physiology and Cognitive Performance. DTIC Document.

Patrick, J., Morgan, P.L., 2010. Approaches to understanding, analysing and developing situation awareness. Theor. Iss. Ergon. Sci. 11, 41–57.

Reason, J., 1990. Human Error. Cambridge University Press, New York.

Sætrevik, B., Eid, J., 2014. The 'similarity index' as an indicator of shared mental models and situation awareness in field studies. J. Cognit. Eng. Decis. Making 8, 119–136.

Salas, E., Dickinson, D.L., Converse, S., Tannenbaum, S.I., 1992. Toward an understanding of team performance and training. In: Swezay, R.W., Salas, E. (Eds.), Teams: Their Training and Performance. Ablex, Norwood, NJ, pp. 3–29.

Salas, E., Rosen, M.A., Burke, C.S., Nicholson, D., Howse, W.R., 2007. Markers for enhancing team cognition in complex environments: the power of team performance diagnosis. Aviat. Space Environ. Med. 78, B77–B85.

Salas, E., Cooke, N.J., Rosen, M.A., 2008. On teams, teamwork, and team performance: discoveries and developments. Hum. Factors: J. Hum. Factors Ergon. Soc. 50, 540–547.

Salmon, P.M., Stanton, N.A., Walker, G., Green, D., 2006. Situation awareness measurement: a review of applicability for C4i environments. Appl. Ergon. 37, 225–238.

Saner, L.D., Bolstad, C.A., Gonzalez, C., Cuevas, H.M., 2009. Measuring and predicting shared situation awareness in teams. J. Cognit. Eng. Decis. Making 3, 280.

Saus, E.-R., Johnsen, B.H., Eid, J., Thayer, J.F., 2012. Who benefits from simulator training: personality and heart rate variability in relation to situation awareness during navigation training. Comput. Hum. Behav. 28, 1262–1268.

Taylor, R.M., 1990. Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design. The Situational Awareness in Aerospace Operations AGARDCP478.

Thayer, J.F., Hansen, A.L., Saus, E.-R., Johnsen, B.H., 2009. Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. Ann. Behav. Med. 37, 141–153.

Thayer, J.F., Åhs, F., Fredrikson, M., Sollers III, J.J., Wager, T.D., 2012. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neurosci. Biobehav. Rev. 36, 747–756.

Wood, R.E., 1986. Task complexity: definition of the construct. Organ. Behav. Hum. Decis. Process. 37, 60–82.

Woods, D.D., Sarter, N.B., 2010. Capturing the dynamics of attention control from individual to distributed systems: the shape of models to come. Theor. Iss. Ergon. Sci. 11, 7–28.