The shared priorities measure as a way of assessing team strategic awareness – a bridge between self-assessment and the deep blue sea of field recordings

Peter Berggren
FOI, Sweden
peter.berggren@foi.se

Björn JE Johansson
FOI, Sweden
bjorn.j.e.johansson@foi.se

Nicoletta Baroutsi
FHS, Sweden
nicoletta.baroutsi@fhs.se

Nils Dahlbäck
LiU, Sweden
nils.dahlback@liu.se

ABSTRACT

Objective, easy to use, easy to comprehend, high face-validity assessment methods for measuring shared awareness in teams are hard to find. This paper describes an experiment where a new measure called Shared Priorities, which is based on ranking of self-generated strategic items, is tested. Trained teams were compared to non-trained teams in a dynamic problem-solving task in terms of performance and shared awareness. The shared priorities measure was used alongside other, well-documented measures of team awareness based on self-rating. The results show that the Shared Priorities measure correlate with performance and could also distinguish between trained and non-trained teams. However, the Shared Priorities measure did not correlate with the other team measures, suggesting that it captures a different quality of team work than the self-rating measures. Further, the shared priorities measure was found to be easily administered and gained a high user acceptance.

Author Keywords

Shared priorities; CARS; team; microworld

INTRODUCTION

Teams are, in most organizations, the basic building blocks for achieving more complicated tasks and goals. Teams are defined as groups of people, whom all take certain roles and conduct specialized tasks aiming at reaching a common goal [36, 32]. Further, team member's actions take place within the same time-frame and are interrelated and interdependent [6]. The way teams perform, and the underlying factors shaping that performance is an area that has gained ever increasing interest during the last decades, especially since the introduction of networked computing that enables team to jointly work on problems, both as co-located and distributed teams [18, 1, 2, 24, 28, 22].

Teams, and what make them effective, have become a major concern for practitioners and researchers. In many cases, findings from group research have been applied on teams [28, 14], but there is also a vast body of research on teams that have emerged (see for example [43, 32] for a comprehensive overview). A large part of this research has focused on investigating the importance of team "awareness" and/or "sharedness" or simply "team cognition", especially in research focused on teams in dynamic control tasks. This is motivated by the idea that teams, just as individuals, must be able to sense, understand and act in order to remain in control of a situation (perform within an acceptable performance envelope, [21, 22]). What mostly is studied (measured) is the "sense and understand" part, in many cases based on the concept of "situation awareness". Situation awareness, as defined by Endsley [13], is a commonly used construct that, if we put it a bit simplistic, is based on the premise that a person who has a good understanding of a situation has a good basis for handling the same situation (this is not necessarily true, but many studies depart from this line of argumentation). The situation awareness concept, has since then been expanded to teams. However, this is not unproblematic. Team awareness or team sharedness comprise more than a mere understanding of what is going on in the surrounding world. Rather, it involves the team members' understanding of the different roles in the team, the assumptions of others' knowledge of other team members' understanding of the situation and even interpersonal relationships. As pointed out by Salas & Cannon-Bowers [40], what's and how's of assessing team performance and awareness or other aspects of teamwork is diverse and lacks consensus.

Wildman et al. [43] has written a comprehensive overview of existing approaches to measuring what could be called "shared awareness" in team. As pointed out by Wildman et al. [43]: "A vast, and often confusing, array of team cognition measurement approaches exist, ranging from simple to complex, from quick to time consuming, and covering a wide range of constructs." (p.1). The methodological approaches used vary greatly, ranging from purely descriptive approaches, such as ethnographic studies, discourse analysis or video analysis, to experimental studies aiming at creating indicators of team performance with predictive power. Wildman et al. [43] (p. 6) have identified six primary data collection methods used in team research: "We have categorized the data collection sources used within the broader team cognition literature into six primary types: (1) interview transcripts, (2) communication transcripts, (3) video records of behavior, (4) direct observations of behavior, (5) self-reported perceptions of team cognition, and (6) self-reported individual knowledge". Wildman et al. [43] discuss the pros and cons of the different methods in relation to what kind of conclusions that can be drawn from the different types of data that are collected. There is no doubt that video- or audio recordings will provide a very rich material and that such methods are possible to apply in a way that does not have to intrude or disturb participants in study. However, the drawback is that such data is very time-demanding to analyze. Only preparation for analysis, transcription, takes at least three to four times longer than the actual recording [43]. Direct observation suffers from the limitations of the observer, i.e. there is a risk that the observer misses important events, that the observer misinterpret what is going on, or that the observer simply is biased in his/her way of observing and taking notes. Direct observation also demands a certain level of domain knowledge, otherwise it will be difficult for the observer to make sense of what he/she is seeing [25, 12, 34]. Self-reported measures are very common in team research, both concerning perceptions of team performance and as individual measures. While possibly giving access to the "inner workings" of a team, they are perceptions, meaning that they do not necessarily represent a good account of what actually is happening.

From a team perspective, self-reporting measures are founded on individual points-of-view, forcing the researcher to apply different kinds of statistical approaches to aggregate a "team measure", usually in the form of average or median calculations of scale-rating replies. These challenges have been discussed by several authors [41, 39, 19]. Depending on the type of administration, such measures are also more or less intrusive [43]. Dynamic measures, which may be repeated several times during for example a training session, will most likely effect the way teams work and how they experience and think about the current situation. "In other words, by measuring team cognition, there is a strong possibility that team cognition is being changed." [43] (p. 24) (cf. metacognition, [20]).

Also, most quantitative measures, such as SAGAT [27], DATMA [31], CARS [33] etc., are explicitly based on theoretical constructs, meaning that a top-down perspective is applied rather than a data-driven such as in the case of the more qualitative approaches. Such models assume that there are some general phenomena that apply to all situations, or rather, unless the underlying model is relevant, the measure most likely lacks construct validity [9]. There are thus pros and cons related to all kinds of data collection methods. On the extremes, we have video recordings and the cumbersome data analysis that results from the method and on the other end simple self-report measures that are easily managed but reduces team interaction and complex situations to pin-pointed quantitative measures based on theoretical constructs.

How should one then bridge the depth between the deep blue sea of purely data-driven methods (which describes team interaction and process in a detailed way) and self-report measures (which may predict performance, but tells us little about team interaction and actual beliefs of team members)? Our view is that "team awareness" emerges from the interactions of the team (cf Cooke, et al. [10]) in relation to reality, and also as a consequence of team (group) formation [42]. From this point of view, there is no "shared situational awareness" in the sense that all team members have the same understanding of the current situation. Rather, each team member has his/her keyhole view of events, which is based on the individual's competence, organization of work, technological constraints [22], as well as earlier perceptions [37]. From that point of view, the idea of "shared situational awareness" becomes less relevant as it actually would be contra-productive to strive towards a completely shared view of ongoing events. Instead, a well-functioning team will act as a joint system that serves the purpose of achieving a common goal, something that does not demand identical situation awareness on the individual level. What then should be of interest for a scientist wanting to understand if a team is functioning well is to ask whether the team members are striving towards the same goal (or goals)? This is similar to the concept of "Strategic Consensus" as described by for example Kellermans et al. [29].

Strategic consensus has mostly been applied to top-level management in companies and there have been some difficulties in establishing a clear relation between the measure and performance, most likely because it has been administered on management teams with different levels of maturity and also in settings signified by low experimental control [29]. Strategic consensus is measured, either by asking participants about the about how much agreement there is within the organization, or by asking individuals to respond to questions regarding a specific strategic approach.

What we suggest is a measure that can capture whether a team has a common view of goals in a given situation, which can be administered easily and analyzed without too much effort. Generally, the researcher aiming at investigating the relation between measures of team shared awareness and performance for teams involved in solving dynamic, complex tasks are faced with two problems: Firstly, the researcher must find a task that has a reasonable level of difficulty but at the same time can provide a performance measure that makes sense. Secondly, the researcher must be able to present a measure of team awareness/cognition that can be correlated with the performance measure. That measure should also be valid, reasonably easy to administer, and acceptable for the participants in that study, i.e. have face validity. Below, we will argue for a new measure called "Shared Priorities", which have some of these characteristics. We will also provide an empirical example where the Shared Priorities measure is compared to some well-established measures of team awareness.

Shared priorities

The Shared Priorities measure is based on the idea that team members rank order items that are directly related to their current work situation. Typically, five items are used. A team member can usually perform this task within 2 minutes. The researcher then calculates Kendall's measure of concordance [30], i.e. to what degree the team members have ranked the items in a similar way. The way to generate items has varied between different studies. Berggren et al. [5] used predefined items in a study on battle tank officers' perception of what was important for the tank crew to perform. In that study, all team members showed the same appreciation of the situation regardless of performance. Later, Berggren and Johansson [4] tested a different approach where the team members generated the items to be ordered. This simplified and improved the shared priorities measure since the items were generated faster, were more closely related to the situation, had higher face-validity, and also needed less preparation time. This study showed a significant effect of the shared priorities measure. Prytz et al. [38] carried out a follow-up study with teams using students as participants and team member generated items. This design showed no main effect of shared priorities. That the shared priorities results from study [38] differ from the former study [4] is probably explained by the fact that student teams were less cohesive than the professional teams since they had no former experience of each other and the task. In the study reported in this paper, this drawback has been eliminated by training half of the participating teams for a longer period (see below). Also, in order to gain an understanding of the validity of the shared priorities measure, we compare the outcome of shared priorities with the commonly used CARS and DATMA measures by means of correlation to performance.

The C3 Fire Microworld

In order to have a controllable, but yet relevant environment, we have chosen to utilize a microworld called C3Fire in this study. Microworlds have been used for several decades and typically present the participants in a study with a situation that is complex, dynamic and opaque [7, 11,8, 23, 35, 15, 26]. At the same time, all conditions in the microworld can be controlled by the researcher and a number of quantifiable performance measures can be retrieved. The microworld thus offer a setting that is experienced as dynamic and complex by the participants in the study, but yet offers a controllable and tractable tool for the researcher.

C3Fire is a microworld that can be used by either an individual or a group of people collaborating with the goal to extinguish a simulated forest fire [16, 17]. The collaboration can be supported by different means of communication and it is possible to configure the simulation so that a dependency is created between the different members. The participants in a C3Fire run typically take the roles of chiefs over a number of vehicles, such as fire engines, water trucks or fuel trucks. They must then coordinate their efforts in order to fight the forest fire as efficiently as possible. The fire engines must fight the forest fire, but they depend on the water trucks to provide them with water for extinguishing the fire. Both fire engines and water trucks depend on the fuel trucks to keep moving. The interactions with the simulated, sub-ordinate units are conducted through a geographical information system, which is an interactive electronic map (see Figure 1).

C3Fire is highly configurable, making it possible for the experimental leader to design the flow of information from units in the simulation to different roles in the fire-fighting organization (the experiment participants), as well as information flow between the participants in the experiment.

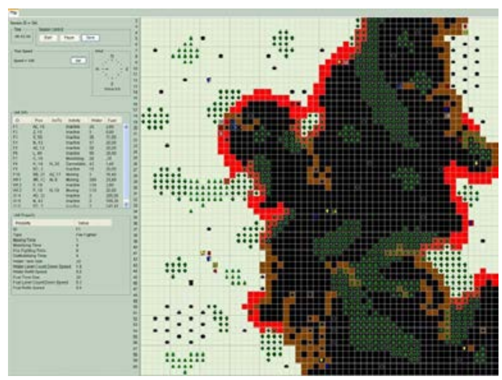

Figure 1. Fire user interface. Black areas are burned down, red areas are burning and brown represent areas where the fire has been successfully extinguished.

PURPOSE/HYPOTHESIS

The main hypotheses tested in this experiment was if 1.) the shared priorities measure can be correlated with performance, if 2.) the shared priorities measure correlates with other, well-established measures of "team awareness", and 3.) if the shared priorities measure could distinguish between trained and non-trained teams. The first two hypothesizes aim to test if the shared priorities measure as a construct is connected to the outcome of the team process, and if it somehow is related to other measures that previously have been connected to performance. The last hypothesis is based on the idea that only "mature" teams are able to understand their own activity in relation to the ongoing events in the environment and their respective roles in the team. Non-trained teams should reasonably be poor at this, while experienced teams should be able to do so. Also, self-rating scores such as CARS and DATMA (see below), should not necessarily be able to catch this, as the underlying construct in those measures never intended to do so.

METHOD

A split plot design was used, team type (trained teams vs non-trained teams) X sensor range (full view scenario vs limited view scenario). Scenario order was balanced over runs. Full view meant that the participants could see what other team members' units were doing and how the fire was developing, regardless of distance to the participants units. Limited view meant that a participant could only see what was happening in cells adjacent to his/her units. This included both units and fire development. To see how a fire is spreading a participant need to have a unit next to it, demanding a more active use of the resources as recon units.

Participants

Twelve teams with three members in each team participated, six trained and six non-trained teams. There were 28 men and 8 women. The mean age of the participants was 28.9 years. There was no significant difference between trained and non-trained teams regarding age or gaming experience. The trained teams were trained for 10 session that included simulator runs and after action discussions. Each training session comprised of a 25 minutes simulator run and about 35 minutes of after action discussion including answering a questionnaire. For details, see Baroutsi et al. [3]. The non-trained teams received a brief introduction including simulator runs and organizational structure (22 minutes). When acting in the scenario, each member of the team chose a role to play: Fire Chief, Water Chief, or Fuel Chief. Each role controlled multiple vehicles. Fire Chief controlled four regular fire trucks, and two fast moving fire trucks (with smaller water tanks). Water Chief controlled three water trucks and two regular fire trucks. Fuel Chief controlled three fuel trucks and two regular fire trucks. The configuration forces the participants to coordinate actions within the team.

Equipment

Four computers were used to run the C3 Fire microworld simulation, one for each participant in the team and one for administering the microworld.

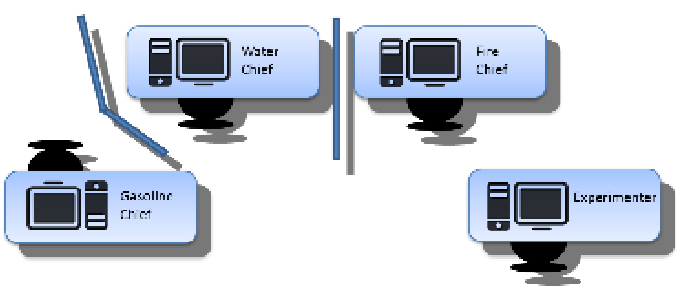

Figure 2. Physical organization of the experiment.

The team members were placed in cubicles so that they would not be able to directly observe each other's screens, thus forcing them to collaborate in order to keep track of each other's activities see Figure 2.

Measures

Dependent variables were simulation performance, CARS [33], mutual awareness [31], and shared priorities.

Performance in the simulation was calculated from the number of burned out cells in the C3 Fire microworld. Different cells had different values depending on the type of object in it. For example, houses were weighted higher than wooded areas, which in turn were weighted higher than grass fields. Each team's score was divided with a baseline score and then subtracted from 1, resulting in a value between 0 and 1 where 0 indicated poor performance and 1 indicated perfect performance.

The CARS-measure included eight questions. For every question the participant is asked to rate themselves on a 4-point scale, adopted from Prytz et al. [38]. The CARS questionnaire was analyzed by calculating the mean for all CARS-questions for each team and scenario.

The DATMA mutual awareness instrument consisted of three parts: team workload awareness, task awareness, and teamwork assessment. DATMA teamwork assessment is rated on a 7-graded scale and DATMA team workload awareness is rated on a 20-graded scale. The modified Task Awareness questionnaire [38], was altered reducing the 9 categories to 7. This is because it is difficult to have opinions about someone else's intentions. For this paper, only team level measures were used. That is, the team workload was calculated using the DATMA team workload awareness questions on team level, excluding the team performance question.

For the shared priorities measure, all team members were asked to individually generate a list of five items that were important for the team to perform in the current situation (target lists). The next step was then to let the team members rank each other's scrambled target lists. The three rank ordered sets of lists could then be analyzed using Kendall's measure of concordance [30], giving a value between 0 and 1. An example of a generated target list is seen below. The list is ordered from first to fifth.

1 Stop the fire from reaching houses and schools around AD55.
2 Refill the units fighting the fire around V54.
3 Close down the fire around C49 and P49.
4 Refill units around Q49.
5 Move the units to the middle of the map.

Procedure

After having received the desired amount of training (see above), an experiment session was performed. The participants could choose among themselves what role to operate: Fire Chief, Water Chief, or Fuel Chief. The participants were divided by screens for viewing obscurity reasons. The teams took part in two simulator runs in C3Fire, each taking 25 minutes. After each run was finished the participants responded to the Shared Priorities instrument, mutual awareness, and CARS. Log-files were automatically saved by the C3Fire simulation server.

RESULTS

All dependent measures were analyzed using a 2 (team type) x 2 (sensor range) ANOVA. The results are compiled in Table 1. There was a significant main effect of sensor range for all dependent measures except shared priorities. The teams performed better in the full view condition. Also, the teams rated themselves higher on CARS, experienced lower workload on the DATMA Team workload instrument, and scored higher on both DATMA Task awareness and Teamwork assessment in the full view condition. For the shared priorities measure there was no main effect of sensor range.

For several of the dependent measures there was a main effect of team type (simulation performance, shared priorities, and DATMA Task Awareness), for workload $p=0.073$. Trained teams performed better than non-trained teams on simulation performance. Trained teams also rated themselves higher on CARS, experienced lower workload on the DATMA Team workload instrument, and scored higher on DATMA Teamwork assessment than non-trained teams

All dependent measures were correlated over both conditions. The correlation matrix can be seen in Table 2. The shared priorities measure correlates significantly with simulation performance, and there is a significant negative correlation with workload. However, shared priorities does not correlate with CARS, nor with task awareness or teamwork assessment. CARS correlates significantly with simulation performance, task awareness, and teamwork assessment. There is also a significant negative correlation with workload.

DISCUSSION

Returning to the hypotheses proposed above; shared priorities measure correlates with performance (hypothesis 1 not rejected). This means that shared priorities, which is not a self-rating measure or a subjective measure of overall team performance, is related to actual performance of the team. The shared priorities measure does not correlate with other, well-established measures (hypothesis 2 rejected). Since there is a correlation with performance, the shared priorities measure reflects other aspects of team cognition than do CARS and DATMA (except for DATMA team workload). Trained teams score better on shared priorities than non-trained teams (hypothesis 3 not rejected), while the other measures do. This suggests that the shared priorities measure captures other aspects of team "sharedness" than the other measures.

Further, all the measures used in this study are more or less intrusive. However, shared priorities could be administered fairly quickly, is easy to analyze and relates directly to the situation in which the participants are at the moment. The other measures on the other hand tries to capture more general aspects of team/situation awareness, but they never demand the participants of the team to actually manifest any of their "cognition" in terms of their actual understanding of the team or the situation, i.e. there is no "objective" data regarding the sharedness of either understanding of what is happening, what will happen or what should be done currently or in the future. Shared priorities at least forces the participants to state what they believe are important to achieve in the current situation. A possible methodological problem is that using a tool like shared priorities may, in itself, alter the "team cognition" as it force/encourage the team members to reflect upon their own activity (cf. metacognition [20]). It should not, however, increase "sharedness" as the team members do not have the possibility to coordinate their ranking of the items. However, just seeing the items may have an impact on behavior. This could also be the case for the other measures, but perhaps not as directly as in the case of shared priorities.

Another question that remains unanswered is how many members a team can have and still be examined with shared priorities? In theory, there is no upper limit, but the larger the team, the more unlikely it is that the ranking items will make sense to the participants as the members of a large team typically becomes more diverse in terms of roles, goals and views of the surrounding world (assuming that a team member typically has his/her "keyhole" perception of what is happening). Most methods approaching the measure of team SA suffer from this drawback, namely that they depart from either self-rating on different scales, or expert ratings of team members. The problem with this approach is that there is a risk that the ratings never actually reflect the actual team SA, but rather the feeling of cohesiveness or possibility of success in a team. The fact that such measures correlate with performance could stem from the fact that the participants in the study understands that they are performing well, which not necessarily must be an indication of "Crew Awareness" or "Mutual Awareness". Also, such measures are constructed in such a way that a team may get fairly high scores on the tests even if the team process actually is suboptimal, while a measure such as Shared Priorities will be unmerciful towards team that actually are unable to achieve a basic level of sharedness in terms of their understanding of current goals. On the other hand, Shared Priorities will most likely not be applicable to immature teams, since it is unreasonable to think that such teams should be able to create a shared understanding. The amount of training a team needs in order to score in a decent way on the Shared Priorities measure is still an open question.

CONCLUSION

In this paper, a microworld study was conducted to investigate if Shared Priorities could be correlated with performance, if it could be correlated with other, established measures of team situation awareness/mutual understanding and if trained teams would score better than non-trained teams on the Shared Priorities measure. Shared Priorities was found to correlate with performance. No correlation was found with the other measures used in the study. Trained teams did score higher on the Shared Priorities measure than non-trained teams. This suggests that the shared priorities measure captures other aspects of team “sharedness” than established measures. Specifically, it focuses on the agreement on objectives/goals rather than perceptions of the current situation. Implicitly, this relates to level three SA, which concerns “projection of future states”, but it goes beyond this as the items describe what is to be achieved rather than what is going to happen. Intent is thus an important aspect of the shared priorities measure compared to the definition of SA as defined by Endsley.

A vast range of measures of “team cognition”, “team awareness”, “team situation awareness” and similar notions exist. In general, the methods associated with these measure either generate vast amounts of data such as field recordings (video, audio) or are based on theoretical constructs and are administered in the form of self-rating scales. The former, although being mostly non-intrusive, suffers from the fact that it is very cumbersome, and expertise-demanding to analyze. The latter often fails to capture team interactions, and, most importantly, does not capture “objective” manifestations of team sharedness. Some measures are based only on outcome parameters or specific tasks, but such measures must be tailored to the exact situation in which the participants work, thus demanding a lot or preparation and expertise.

In comparison to other measures of team cognition, the Shared Priorities measure demand little preparation, demand no expertise in the domain in which the participants take part on part of the researcher and is fairly straight-forward to analyze.

REFERENCES

1. Artman, H. and Waern, Y. Creation and Loss of Cognitive Empathy at an Emergency Control Center. In (Ed.) Y. Waern. *Cooperative Process Management – Cognition and Information Technology*. Francis & Taylor: London, 1998.
2. Artman, H. Team situation assessment and information distribution. *Ergonomics 43*,8, (2000), 1111-1128.
3. Baroutsi, N., Berggren, P., Nählinder, S. and Johansson, B. *Training teams to collaborative as cohesive units* (Scientific report No. FOI-R--3830--SE), FOI, 2013.
4. Berggren, P. and Johansson, B. Developing an instrument for measuring shared understanding. In *Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management: Defining Crisis Management 3.0*. Seattle, WA. May 2-5 (2010).
5. Berggren, P., Svensson, J. and Hörberg, U. *Mätning av gemensam lägesbild vid ledning på stridsteknisk och taktisk nivå - Studie genomförd på TCCS* (Användarrapport No. FOI-R--2647--SE), FOI, 2008.
6. Brannick, M.T., and Prince, C. An Overview of Team Performance Measurement. In (Eds.) M.T. Brannick, E. Salas, and C. Prince, *Team performance Assessment and Measurement*. Lawrence Erlbaum Associates, London, 1997.
7. Brehmer, B. and Dörner, D. Experiments With Computer-Simulated Microworlds: Escaping Both the Narrow Straits of the Laboratory and the Deep Blue Sea of the Field Study. *Computers in Human Behaviour, 9*, (1993), 171-184.
8. Brehmer, B. Dynamic decision making: Human control of complex systems. *Acta Psychologica*. *81*, (1992), 211-241.
9. Cook, T.D. and Campbell, D.T. *Quasi-experimentation: Design and anlysis issues for field settings*. Houghton Mifflin Company, Boston, 1979.
10. Cooke, N.J., Gorman, J.C., Myers, C.W. and Duran, J.L. Interactive team cognition. *Cognitive Science 37,* 2, (2013), 255–85.
11. Dörner, D., Kreuzig, H.W., Reiter, F. and Städel, T. *Lohausen. Von Umgang mit Unbestimmheit und Komplexität*. Huber, Bern, 1983.
12. Drury, C.G. Methods for direct observation of performance. In (Eds.) J.R. Wilson and E.N. Corlett. *Evaluation of Human Work*. Taylor and Francis, London, 1992.
13. Endsley, M.R. Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors 37*, 1, (1995), 32–64.
14. Fullagar, C.J., and Egleston, D.O. Norming and Performing : Using Microworlds to Understand the Relationship Between Team Cohesiveness and Performance. *Journal of Applied Social Psychology*, *38, 10*, (2008), 2574–2593.
15. Gonzalez C., Vanyukov, P. and Martin M.K. The use of microworlds to study dynamic decision making. *Computers in Human Behavior 21*, (2005), 273-286.
16. Granlund R, Johansson B, and Persson M. C3Fire a Micro-world for Collaboration Training in the ROLF environment. In *proceedings to SIMS 2001 the 42nd Conference on Simulation and Modelling, Simulation in Theory and Practice,* (2001).
17. Granlund R. *Monitoring Distributed Teamwork Training*. Linköping Studies in Science and Technology, Linköping University Press, Linköping ,2002.
18. Heath, C.C. and Luff, P. Collaboration and control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms. *Computer*

Supported Cooperative Work (CSCW). An International Journal, vol. 1, nos. 1–2, (1992), 69–94.

19. Helton, W.S., Funke, G.J. and Knott, B.A. Measuring Workload in Collaborative Contexts: Trait Versus State Perspectives. *Human Factors, online*, (2013), 1–11.
20. Hinsz, V.B. Metacognition and mental models in groups: An illustration with metamemory of group recognition memory. In (Eds.) E. Salas and S. M. Fiore, *Team cognition: Understanding the factors that drive process and performance*, American Psychological Association, Washington, DC, 2004, 33–58.
21. Hollnagel, E. *Cognitive reliability and error analysis method — CREAM*. Elsevier Science, Oxford, 1998.
22. Hollnagel, E., and Woods, D.D. *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC Press, Boca Raton, FL, 2005.
23. Howie, D.E. and Vincente, K.J. Measures of operator performance in complex, dynamic microworlds: advancing the state of the art, *Ergonomics 41*, 4, (1998), 485-500.
24. Hutchins, E. and Klausen, T. Distributed Cognition in an Airlane Cockpit. In (Eds.) Y. Engeström and D. Middelton. *Cognition and Communication at Work.* Cambridge University Press, Cambridge, 1996.
25. Jersild, A.T. and Meigs, M.F. Direct Observation as a Research Method. Review of Educational Research. *Methods of Research in Education 9*, 5, (1939) 472-482.
26. Jobidon, M.-E., Breton, R., Rousseau, R., and Tremblay, S. Team response to workload transition: The role of team structure. In *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, CA, (2006).
27. Jones, D.G. and Kaber, D.B. Situation awareness measurement and the Situation awareness global assessment technique. In (Eds.) N.A. Stanton, A. Hedge, K. Brookhuis, E. Salas and H. Hendrick, *Handbook of Human Factors and Ergonomics methods*. CRC Press, London, 2005.
28. Jones, P.E. and Roelofsma, P.H.M.P. The potential for social contextual and group biases in team decision-making: Biases, conditions and psychological mechanisms. *Ergonomics, 43,* 8, (2000), 1129-1152.
29. Kellermanns, F.W., Walter, J., Lechner, C., and Floyd, S.W. The Lack of Consensus About Strategic Consensus: Advancing Theory and Research. *Journal of Management*, *31*, 5, (2005), 719–737.
30. Kendall, M.G. *Rank correlation methods* (Forth ed.).: Charles Griffin & Company Ltd, London, 1975.
31. Macmillan, J., Paley, M.J., Entin, E.B. and Entin, E.E. Questionnaires for Distributed Assessment of Team Mutual Awareness. In (Eds.) N.A. Stanton, A. Hedge, K. Brookhuis, E. Salas and H. Hendricks. *Handbook of Human Factors Methods*, Taylor and Francis, London, 2005.
32. Mathieu, J., Maynard, M.T., Rapp, T. and Gilson, L. Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management 34*, 3, (2008), 410-476.
33. McGuinnes, B. and Foy, L.A. Subjective measure of SA: the Crew Awareness Rating Scale (CARS). *In The Human Performance, Situational Awareness an Automation Conference,* (2000).
34. Miles, M.B. and Huberman, A.M. *Qualitative Data Analysis: An Expanded Sourcebook*. Sage Publishing: London, 1994.
35. Omodei, M.M., Wearing, A.J., McLennan, J., Elliot, G.C. and Clancy, J.M. “More is better?”: Problems of Self-Regulation in Naturalistic Decision Making Settings. In (Eds.) B. Brehmer, H. Montgomery, and R. Lipshitz. *How Professionals make decisions*, Lawrence Erlbaum Associates Inc.: Mahaw, New Jersey, 2004.
36. Orasanu, J. and Salas, E. Team Decision Making in Complex Environments. In (Eds.) Klein, G.A., Orasanu, J., Calderwood, R. and Zsambook, C.E. *Decision Making in Action*. Ablex Publishing, Norwood, NJ, 1993.
37. Peirce, C. S. *Collected Papers of Charles Sanders Peirce, vols. 1–6, 1931–1935,* In (Eds.) C. Hartshorne & P. Weiss. Harvard University Press, Cambridge, MA, 1958.
38. Prytz, E., Berggren, P. and Johansson, B.J.E. *Performance and shared understanding in mixed C2-systems* (Scientific report No. FOI-R--3155--SE), FOI, 2010.
39. Saetrevik, B. and Eid, J. The “Similarity Index” as an Indicator of Shared Mental Models and Situation Awareness in Field Studies. *Journal of Cognitive Engineering and Decision Making*, (2013), Online first: 1.18.
40. Salas, E. and Cannon-Bowers, J.A. Special issue preface. *Journal of Organizational Behavior*, *22*, (2001), 87–88.
41. Saner, L.D., Bolstad, C.A., Gonzales, C. and Cuevas, H.M. Measuring and Predicting Shared Situation Awareness in Teams. *Journal of Cognitive Engineering and Decision Making 3*, 3, (2009), 280-308.
42. Tuckman, B.W. Developmental Sequence in Small Groups. *Psychological Bulletin 6*,36, (1965), 384-399.
43. Wildman, J.L., Salas, E., and Scott, C.P.R. Measuring Cognition in Teams A Cross-Domain Review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *Online*, (2013), 1–31.

Table 1. Table of ANOVAs performed including means for the different conditions.

Dependent measure	Main effect of sensor range	Mean full view	Mean limited view	Main effect of team type	Mean trained teams	Mean non-trained teams	Interaction effect
Simulation performance	$F(1,10)=29.63$, $p<0.001$	0.60	0.39	$F(1,10)=15.38$, $p<0.005$	0.65	0.34	ns
Shared priorities	ns	0.55	0.43	$F(1,10)=8.44$, $p<0.05$	0.58	0.40	ns
CARS	$F(1,10)=38.20$, $p<0.001$	3.16	2.46	ns	2.94	2.68	$F(1,10)=9.02$, $p<0.05$
DATMA Team Workload	$F(1,10)=13.81$, $p<0.005$	10.78	12.98	$F(1,10)=4.02$, $p=0.073$	10.80	12.96	$F(1,10)=3.57$, $p=0.088$
DATMA Task Awareness	$F(1,10)=10.65$, $p<0.01$	8.68	6.67	$F(1,10)=5.28$, $p<0.05$	9.44	5.90	ns
DATMA Teamwork Assessmen	$F(1,10)=27.82$, $p<0.001$	14.94	11.73	ns	14.33	12.34	$F(1,10)=8.59$, $p<0.05$

Table 2. Correlation matrix for the different dependent measures.

	Simulation performance	Shared priorities	CARS	DATMA Team Workload	DATMA Task Awareness
Shared priorities	,496*				
CARS	,536**	,190			
DATMA Team Workload	-,611**	-,456*	-,644**		
DATMA Task Awareness	,618**	,335	,457*	-,729**	
DATMA Teamwork Assessment	,661**	,199	,908**	-,577**	,373

*. Correlation is significant at the 0.05 level (2-tailed), **. Correlation is significant at the 0.01 level (2-tailed).