Register

Study Information

Title

Provide the working title of your study. It is helpful if this is the same title that you submit for publication of your final manuscript, but it is not a requirement.

The effect of communication type and social networks on shared beliefs in emergency response teams

Authors

The author who submits the preregistration is the recipient of the award money and must also be an author of the published manuscript. Additional authors may be added or removed at any time.

Bjørn Sætrevik, Malin Elisabet Lilletvedt, Annette Falch Lothe

Research Questions

Please list each research question included in this study.

1. Does the type of communication have an effect on 'shared mental models' (SMM) and 'situation awareness' (SA) in emergency response teams?

2. Does the structure of the social network in the emergency response teams have an effect on SMM and SA?

Hypotheses

For each of the research questions listed in the previous section, provide one or multiple specific and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here.

Hypotheses related to the team leader's communication:
H1a: Teams with more 'closed-loop' communication (relative to the amount of 'confirmatory', 'transfer' and 'request' communication) will have higher SMM.
H1b: Teams with more 'confirmatory' communication (relative to the amount of 'transfer' and 'request' communication) will have higher SMM.
H1c: Teams with more 'transfer' communication (relative to the amount of 'request' communication), will have higher SMM.
H2a: Similarly, teams with more closed-loop' communication will have more accurate SA.
H2b: Teams with more 'confirmatory' communication will have more accurate SA.
H2c: Teams with more 'transfer' communication will have more accurate SA.

Hypotheses related to individual team member changes scenario time:
H3a: Individual SMM will increase across scenario time.
H3b: Individual SA accuracy will increase across scenario time.
H4a: Individual communication connectivity will increase over scenario time.
H4b: Individual reliance connectivity will increase over scenario time.

Hypotheses related to the team's social network:
H5a: Teams with more distributed communication networks will have higher SMM.
H5b: Teams with more distributed communication networks will have more accurate SA.
H6a: Teams with more distributed reliance networks will have more accurate SMM.
H6b: Teams with more distributed reliance networks will have more accurate SA.

Sampling Plan

Existing Data

Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation. Please do not hesitate to contact us if you have questions about how to answer this question (prereg@cos.io).

Registration prior to any human observation of the data

Explanation of existing data

If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study. The purpose of this question is to assure that the line between confirmatory and exploratory analysis is clear.

Data collection is underway, where 6 out of a planned sample of 12 data collections have been completed at the time of submission of the pre-registration. The communication observed during the scenarios have been coded by the researchers, but no counting of the frequency of different communication types have been done. The team members' responses to the SMM, SA and communication and reliance network probes have not been examined (apart from verifying that data has indeed been collected).

Data collection procedures

Please describe the process by which you will collect your data. If you are using human subjects, this should include the population from which you obtain subjects, recruitment efforts, payment for participation, how subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion rules), and your study timeline. For studies that don't include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.

We will collect data under 12 scenario exercises in the emergency center of an international hydrocarbon energy company. The scenarios will be planned and executed by our industry partner, with minimal intrusion of our data collection as described below. Data collections 1-6 (for all six teams) will be a fire / explosion scenario, and data collections 7-12 (the same six teams) will be a loss-of-stability scenario. All data will be collected once a week across 14 consecutive weeks (with a two-week break between data collection 6 and 7).

An exercise lasts for approximately 2 hours. During this exercise, the scenario play will be 'frozen' 4 times, once at the beginning of the exercise, and at every following half-hour (due to administration errors, the first data collection has only two freeze points). If a status meeting is ongoing or has been announced at the scheduled time, the freeze point will be postponed until after the meeting. At each freeze point, the researchers will orally ask the team members to stop their work, and to open a link to an online questionnaire that they received by email at the scenario start. The content of the questionnaire can be found under 'Measured variables'. The questionnaire is spread across four pages, and acceptable responses must be filled out on each page in order to proceed in the questionnaire. Filling out the questionnaire takes about 5 minutes. Team members are instructed not to exchange any information during the freeze point. When the questionnaire has been answered, the team members can return to their tasks.

Throughout the scenario, two researchers (second and third author on pre-registration) will follow the person filling the team role of 'operation section chief' (OSC), and will simultaneously classify all the communication that the OSC has with other team members.

Each statement made by the OSC or directed at the OSC will be scored as either request, transfer, confirmed communication or closed-loop communication (see 'Measured variables' for operationalization of these). The coding will be done with pen-and-paper coding sheets and the researchers will use synchronized watches to ensure consistent coding.

Coding scheme for communication.pdf
Freeze point questions (original in Norwegian).pdf
Freeze point questions (English translation).pdf

Sample size

Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?

The participants in the study are all members of the emergency response teams from a hydrocarbon industry company. The teams will be used in their original configurations. There are six teams that each will be tested twice. The teams consist of up to 15 members, but only 9 of the participants are seated together and collaborate directly in the team work, while 3 are positioned elsewhere in the building, and the final 3 are placed at other facilities. We expect highest compliance from the 9 team members that are located together. As these are also the members most likely to be influenced by the team processes in question, only these responses will be entered into the analysis.

For testing H1-H2 and H5-H6, the variables of interest will be analysed on the team level (averaged across members in a team and across the different freeze points in each scenario run), resulting in a sample size of n=12. For testing H3-H4 and various follow-up analyses, the variables of interest will be retained on individual team member level, resulting in a maximum of n=180 (6 teams x 15 members), but is more likely to be below n=109 (6 teams x 9 co-located members).

Sample size rationale

This could include a power analysis or an arbitrary constraint such as time, money, or personnel.

The whole population of the emergency preparedness organization in the industrial company is invited to participate in the experiment. All six teams will participate twice, but the participation of individual members in the team may vary due to availability and compliance. Variations in manning may lead to some team members taking part in one scenario more than once.

Stopping rule

If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.

Not relevant for this study.

Variables

Manipulated variables

Describe all variables you plan to manipulate and the levels or treatment arms of each variable. For observational studies and meta-analyses, simply state that this is not applicable.

We will not manipulate any variables.

no file selected

Measured variables

Describe each variable that you will measure. This will include outcome measures, as well as any predictors or covariates that you will measure. You do not need to include any variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this study.

Two researchers will simultaneously observe the communications that the 'operation section chief' (OSC) has with other team members, and will continuously code it in a pen-and-paper coding scheme. Every oral statement given by or to the OSC will be coded as either 'request', 'information transfer', 'confirmation' or 'closed-loop communication' using the following operational definitions: 'Request' (R) is a statement that seeks information, action or confirmation. 'Transfer' (T) is a statement that shares information with one or more team-members. 'Confirmatory' (C) is a statement that acknowledges that a previous statement or action by another team member has been received or understood. 'Closed-loop' (CL) is a statement that confirms or rejects another team members repetition or interpretation of one's own previous statement. During observation, every statement is noted with the starting clock time, and the breakpoint between conversations will be marked.

'SMM' and 'SA' are calculated based on the answers to questions about the team member's beliefs about the scenario task and the team's work that all team members should be familiar with. The beliefs are measured in an online questionnaire that are sent out as an email at the beginning of the exercise containing separate links for each freeze point (see attached under 'Data collection procedures', probe 1-5). The questionnaire contains questions about (probe 1) what part of the installation that is afflicted (answered by ticking of checkboxes next to 6 different areas, and 'don't know'), (probe 2) about what kind of incident is currently active (answered by ticking of checkboxes next to 15 different types of incident, as well as 'don't know' or 'none of the above'), (probe 3) questions about how many are still on board, how many are missing, injured or deceased (can answer 'none' or enter text or number), (probe 4) about how likely seven different outcomes of the event are (each rated from 'certain' to 'irrelevant'), and finally (probe 5a, 5b and 5c) questions about what the team's current first, second and third priorities are (select one of 13 different options at for each of the three priority levels). 'SMM' will be calculated as the similarity between an individual team member's answer and the answers from the whole team. 'SA' will be calculated as the similarity between an individual team member's answer and the answers from the OSC. The calculations are shown under 'Indices'.

Social network analysis will be measured by asking six questions in the online questionnaire about communication and reliance in the emergency response team under a training scenario (attached under 'Data collection procedures', probe 6 and 7). On probe 6a, 6b and 6c each respondent ranks which of the other team members they 1) communicated most with, 2) communicated second most with, and 3) communicated third most with. On probe 7a, 7b and 7c each respondent ranks which of the other team members they 1) relied most on, 2) relied second most on, and 3) relied third most on in solving their task. On each question, the respondent selects one of the team member roles from a list.

no file selected

Indices

If any measurements are going to be combined into an index (or even a mean), what measures will you use and how will they be combined? Include either a formula or a precise description of your method. If your are using a more complicated statistical method to combine measures (e.g. a factor analysis), you can note that here but describe the exact method in the analysis plan section.

Communication indices:
Indices will be calculated for each data collection to reflect the relative amount of each type communication. For testing H1a and H2a we will calculate an index for the amount of 'closed loop' communication compared to the other types (calculated as the number of instances of CL divided by the sum of instances of R, T and C). For testing H1b and H2b we will calculate an index for the amount of 'confirmatory' communication compared to less iterative types of communication (calculated as the number of instances of C divided by the sum of instances of T and R). For testing H1c and H2c we will calculate an index for amount of 'transfer' communication compared less iterative types of communication (calculated as the number of instances of T divided by the sum of instances of R).

Shared belief indices:
Indices will be calculated from the responses to questions about the team members beliefs about the scenario (probe 1-5, see 'Transformations'), following a previously published approach (Sætrevik & Eid, 2014, Journal of Cognitive Engineering and Decision Making).

The SMM index quantifies the similarity between individual team members' beliefs about the scenario and the beliefs of the other team members. For each response option at each probe we will calculate the difference between the team member's answer and the team's average answer, divided by the highest value given on that response option. These will be averaged to an SMM index for each probe item for each team member at each time point, which can be averaged to the individual's average SMM across all four time points in the scenario, or the team's average SMM at a given time point or across the scenario.

The SA index quantifies the similarity between individual team members' beliefs about the scenario and the team leader's beliefs. For each response option at each probe at each time point we will calculate the difference between the team member's answer and the team leader (OSC), divided by the highest value given on that response option. These will be averaged to an SA index for each probe item for each team member at each time point, which can be averaged to the individual's average SA across all four time points in the scenario, or the team's average SA at a given time point or across the scenario.

See attached file that shows how SMM and SA indexes are calculated using different types of mock-up data.

Social network indices:
Indices will be calculated on the individual team member level that will then be aggregated to the team level indices that will be used in the analyses. The approaches are similar for the 'communication' and the 'reliance' variables.

Individual popularity index: For each team member at each freeze point, the number of other team members that have mentioned this team member as the one communicated most with or relied most on will be counted. Being mentioned as the one most important will be given a weight of 3, mentioned as second most important will be weight of 2, and being the third most important will be given a weight of 1. As the popularity will be influenced by how many team members that respond at a given freeze point, the value will be adjusted by dividing the individual score by the sum of the scores for all team members on that freeze point. The resulting index will correspond to a weighted (adjusted for ranking by other members), tuned (adjusted to team member size) in-degree centrality score, and will be referred to as the individual popularity score.

To get a measure of distribution of the networks, we will use the degree of variation in a team's popularity scores (the sum of squared differences between team members in the popularity index at a given freeze point). A team with more variation (a few members have high popularity scores while others have low scores) is interpreted as a more centralized team network, while a team with less variation (most members have similar popularity scores), is interpreted as a more distributed network.

See attached file that show how team member popularity index and aggregated team distribution scores are calculated for different types of mock-up data.

Example of SNA calculations for communication and reliance.xlsx
Example of SMM and SA index calculations.xlsx

Design Plan

Study type

Please check one of the following statements

Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, "natural experiments," and regression discontinuity designs.

Blinding

Blinding describes who is aware of the experimental manipulations within a study. Mark all that apply.

No blinding is involved in this study.

Study design

Describe your study design. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required. Typical study designs for observation studies include cohort, cross sectional, and case-control studies.

This is a natural experiment in a field setting. The experiment tests the effect of arbitrary variation between teams and individuals in terms of communication (observed relative frequencies of 4 different communication types) and of social networks (team member popularity and network distribution) on the degree of shared beliefs among teams and among team members. Effects are tested both as between subjects (H1-H2, and H5-H6), and within-subjects (H3-H4).

no file selected

Randomization

If you are doing a randomized study, how will you randomize, and at what level?

This study is not randomised.

Analysis Plan

Statistical models

What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions that will be tested and remember that any test not included here must be noted as an exploratory test in your final article.

Prior to the hypothesis testing, we will test some of our underlying assumptions for data collection in this setting. The inter-reliability of communication scoring will be calculated, using Cohen's kappa. We will test the assumption that the OSC's communication becomes less frequent as the scenario progresses (two ANOVAs will test the effect of time on the number of statements and the number of conversations preceding the second, the third and the fourth freeze points). We will test whether the overall number of statements and conversations the OSC is involved in will positively correlate with SMM and SA. Further we will test the assumption that the OSC has higher 'popularity' than the other team roles.

To test H1a, H1b, H1c, H2a, H2b and H2c, six independent multiple regressions will be performed to assess the relationship between the relative amount of communication types (as independent variable) and average SMM or SA for each team (as dependent variable). These analyses will use data averaged across the four freeze points.

To test H3a, H3b, H4a, and H4b, four independent ANOVAs will be performed to test the effect of freeze point time (independent variable), on team member scores for SMM, SA communication popularity, and reliance popularity measured at each time point (dependent variable).

To test H5a, H5b, H6a and H6b, four independent multiple regressions will test the association between the team's degree of distributed communication and reliance (independent variables), and the SMM and SA scores (dependent variables). These analyses will use data from the four freeze points separately.

no file selected

Transformations

If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.

Communication coding transformations:
Statements that are interpreted as 'Request' will be scored as 1, 'Transfer' will be scored as 2, 'Confirmatory' will be scored as 3 and 'Closed-loop' will be scored as 4. After digitising the coding sheets, the two researchers' codes will be compared to establish inter-coder reliability. The initial analysis will only use statements that have been coded by both researchers, and where they both have applied the same code. If inter-rater reliability is low and much data is excluded when using this approach, further analyses may also include statements where the coders have applied adjacent codes (e.g. statements coded as 1 by one researcher and 2 will be used, but not statements coded as 1 by one researcher and 3 by the other).

Shared belief probe response transformations:
For probes with multiple choice questions (probe 1 and 2), the response options that have been ticked by a team member will be scored as 1, while response options unticked by this team member but ticked by other members in the team at this time point will be scored as 0 (response options unticked by the whole team at this time point will be left blank). If a respondent has ticked 'none' (or ticked 'Other' and typed in a string indicating the same) in addition to another response option on probe 1 or 2, the other options will be removed, leaving only 'none' as a response. If a respondent has selected 'don't know' instead of or in addition to another response option, the index will be set to 0 for that probe. For probe 3, the inputted number of personnel for each status category will be used (replacing "none" with 0). For probe 4, the different response options (from "certain" to "irrelevant") will be scored as 1-5 (not possible to leave response options blank on probe 3 and 4). For probe 5, the response options that was selected as the highest team priority will be scored as 3, the second priority will be scored as 2, and the third priority will be scored as 1. Response options not selected by this team member but by other team members at this freeze point will be scored as 0 (and left blank if not selected by other team members at that freeze point). All response options will be left blank for participants that have not answered any probes at a given freeze point (even if they have answered at other freeze points in the scenario). See first sheet of Excel file attached under 'Indices' for example of what the transformed scores can look like.

Social network transformations:
For probe 6 and 7, if a team member states that their own team role as one of the persons they have communicated with or relied on, this response will be removed. If a team member states the same team role twice (e.g. that the same role is said to be both the one most and second most communicated with), the highest ranking will be used while removing the lower rank. If higher ranked responses are missing in probe 6 or 7 (either as input or as result of the previously mentioned transformations) while there are lower ranked responses, the lower ranked responses will be moved up to fill the place of the missing response.

Follow-up analyses

If not specified previously, will you be conducting any confirmatory analyses to follow up on effects in your statistical model, such as subgroup analyses, pairwise or complex contrasts, or follow-up tests from interactions? Remember that any analyses not specified in this research plan must be noted as exploratory.

As the first data collection deviated from the remaining eleven (only two freeze points were used instead of four), the analyses will be tested again while excluding this data collection (and any subsequent data collections that end up missing freeze points).

Testing effects of communication types at each freeze point:
If the analyses of H1 or H2 are significant, these will be followed-up by testing the effects of communication types preceding each time point on the SMM and SA scores at the corresponding time point.

Testing the effects of network distribution across scenario:
If the analyses of H5 or H6 are significant, they will be followed-up by testing the effect on scores averaged for the whole scenario, using arithmetic averages from the four freeze points of scores for communication distribution, reliance distribution, SMM and SA.

Testing the effects of individual team member popularity:
If the planned analyses show that degree of team network distribution for communication and/or reliance is significant (H5 or H6), this will be followed up by analyses testing the effects of individual team member communication and reliance popularity on individual SMM and SA.

Inference criteria

What criteria will you use to make inferences? Please describe the information you'll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?

Our hypotheses are supported when the p-value is less than .05, using one-tailed tests due to our directed hypotheses.

We will use Bonferroni correction to counteract the problem of multiple comparisons. As the different classes of hypotheses (H1-H2, H3-H4, and H5-H6) use different independent variables, the corrections will be done within each class, correcting for 6, 4 and 4 analyses, respectively.

Data exclusion

How will you determine which data points or samples (if any) to exclude from your analyses? How will outliers be handled?

In the coding and cleaning of communication type, meetings or telephone calls that the OSC has with people not in the team will not be scored. Statements that are interpreted as not being related to the scenario play will be excluded from the analysis. Statements that have not been noted by both of the two observing researchers (with an onset of not more than 10 seconds apart) will be excluded from the analysis. Statements that have been coded differently by the two researchers will be excluded.

Answer to the online freeze point questionnaire that arrives more than 10 minutes later than the first response will be excluded from the analysis.

Missing data

How will you deal with incomplete or missing data?

The freeze point questionnaires require valid responses on each question in order to proceed in the questionnaire. All data collected with the questionnaire will be used in the analysis, with the transformations described above.

Due to compliance and administrative errors, all four freeze points may not be completed for all data collections (the first data collection had only two freeze points). Analyses that use scenario averages (H1, H2, H5 and H6), will calculate averages from the remaining freeze points.

Exploratory analysis

If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time.

If inter-rater reliability of communication coding is low and much data is excluded when coding into four categories and excluding data with mismatch between the two coders, other approaches will be tested. Exploratory analyses may also include statements where the coders have applied adjacent codes (e.g. statements coded as 1 by one researcher and 2 will be used, but not statements coded as 1 by one researcher and 3 by the other). We may test other approaches, such as collapsing the adjacent categories T and R and/or collapsing the categories C and CL.

The current data collection uses two different types of scenarios, testing all teams once on each scenario. While the analyses mentioned above disregards this variation, exploratory analyses may test for differential effects between the two scenarios.

Analyses of team member level data (H3, H4 and various follow-up analyses) will be explored with additional analyses where 'team membership' is added as a independent variable, to see how much of the effect on dependent variable can be accounted for by team factors (shared experience, leadership, etc).

The analyses of social network data may also explore other measures of team-level network distribution, e.g. density measures (an index of the number of connections made between team members relative to the number of possible connections).

Scripts

Upload an analysis script with clear comments

This optional step is helpful in order to create a process that is completely transparent and increase the likelihood that your analysis can be replicated. We recommend that you run the code on a simulated dataset in order to check that it will run without errors.

no file selected

Other

Other

If there is any additional information that you feel needs to be included in your preregistration, please enter it here.

No additional information.