A Preliminary Psychometric Validation of the Safety Perception Index on Norwegian Vessels (SPIN-V)

Asbjørn Lein Aalberg
SINTEF Digital, Norway. E-mail: asbjorn.lein.aalberg@sintef.no

Rolf Johan Bye
SINTEF Digital. E-mail: rolf.johan.bye@sintef.no

Anastasiia Tkalich
Norwegian University of Science & Technology, Norway, E-mail: nsondor@gmail.com

Leif Inge K. Sørskår
Safetec Nordic AS, Norway. E-mail: leif.inge.sorskar@safetec.no

Vegar Berntsen
Norwegian Maritime Authority, Norway. E-mail: vegar.berntsen@sdir.no

The objective of the study was to examine the psychometric properties of a questionnaire developed for The Norwegian Maritime Authority (NMA), the Safety Perception Index on Norwegian Vessels (SPIN-V), and how different aspects of safety climate and work climate relate to safety outcomes. The aim of the questionnaire was to measure perceptions, attitudes and behavior related to safety and work environment on Norwegian vessels. The results were based on a survey carried out on seafarers in 2019 (N = 8095, 25 % response rate). A combination of exploratory factor analysis, confirmatory factor analysis and partial least square structural equation modelling (PLS-SEM) resulted in a safety climate and work condition structure of seven dimensions: safe work practices, management´s safety involvement, Procedure compliance, competence and training, manning, job demands and safety system. Regression analysis indicated that the identified dimensions were significantly related to subjective safety perception and whether the seafarer had been involved in an accident or near-miss the last year. All of the dimensions predicted subjective safety perception. Self-reported accident involvement was associated with three dimensions: safe work practices, competence and training and Procedure compliance. Self-reported near-miss involvement was related to all dimensions except manning and management´s safety involvement. Violation of procedures was the major predictor of both accident and near-miss involvement. The study provides a basis for further validation and research with the SPIN-V data.

Keywords: maritime safety, safety perception, psychometric, Norwegian vessels.

1. Introduction

There are relatively few published scientific studies based on questionnaires for measuring human perceptions within the maritime domain compared to other industries (Nævestad, 2017). These publications are mainly conducted in order to measure the concepts of safety culture, safety climate or working condition. Main contributors that have published in scientific journals includes e.g. Ek et al. (2000), Ek and Akselsson (2005), Håvold (2005, 2007, 2010a, 2010b, 2015), Darbraet al. (2007), Lu and Tsai (2008), Shang and Lu (2009), Håvold and Nesset (2009), Oltedal and Wadsworth (2010), Håvold et al. (2011), Ek et al. (2014), Berheim et al. (2015), Aslan et al. (2016), and Nævestad (2017). The previous studies have been based on different questionnaires aimed at different segments of the maritime industry. Among the relatively limited number of journal publications based on questionnaires in the maritime sector, few address the question of how good the measurement tool is in terms of validity.

The originality of this study is that it is based on a survey that covers the entirely Norwegian fleet, including the belonging two registers of The Norwegian Ordinary Ship Register (NOR) and The Norwegian International Ship Register (NIS). Further, the survey is designed in order to be used as indicators for policy making and implementation of measures by the Norwegian Maritime Authority (NMA). The number of

Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference
Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio
Published by Research Publishing, Singapore.
ISBN: 978-981-14-8593-0; doi:10.3850/978-981-14-8593-0

respondents in this study has also been high compared to previous studies. The largest number of respondents in a corresponding published maritime survey has so far been about 2500 (Håvold and Nesset, 2009).

Questionnaires measuring employees' perceptions in relation to safety are not uncommon within other industries. For example, in the Norwegian petroleum industry, the Petroleum Safety Authority (PSA) bi-annually issues a questionnaire which invites all employees working on installations to participate. It is used as part of monitoring risk level, input to risk-based supervision, and to shed light on specific issues. One of the goals is also to trend workers' perceptions over time. The questionnaire has been used in several research settings, like Tharaldsen et al. (2008), and Kvalheim and Dahl (2016).

Several criteria (psychometric properties) exist for evaluating the dimensionality and "goodness" of an instrument. In general, the instrument should have sufficient reliability and validity. Support for reliability and validity is strengthened by applying different psychometric approaches, and also by testing and re-testing on different subgroups of the population, time periods and settings.

1.3 *Validity of self-reporting measures*

A central question concerned with tests of humans self-reported perceptions, behavior, attitudes or emotions, are whether what you are measuring actually represents the latent construct intended to measure (Murphy and Davidshofer, 2005). The construct validity may be defined as whether the observed correlations between constructs in the tests are in accordance with current theoretical notions of the constructs (Cronbach and Meehl, 1955). A sub-set of validity aspects are discriminant and convergent validity. In this line of thought, validity is ensured if the constructs are non-related, low or moderately related to constructs they are assumed to be different from, and whether the constructs are in fact correlated with constructs that they are (theoretically) related to (Taherdoost, 2016). Additionally, the content validity, which concerns whether the aspects measured indeed covers the aspects of the theoretical descriptions of the phenomenon, is useful particularly in the domain of questionnaires. More recent, predictive validity (a type of criterion validity) has been put forth as part of a more holistic view of the question of validity (Engvik, 1999). Construct validity can also be seen as a function of the "sum" of all validity types.

Essentially, reliability in this setting can be described as the consistency of the measurement method across time, assessments, assessors or items when the latent phenomenon remains unchanged. The reliability of the scales involved in psychometric measures are often assessed by investigating the consistency across sub-items within the same scale (Field, 2006), which is a common way of assessing reliability of a cross-sectional questionnaire like in the present study.

Construct validity and reliability in measurement/test theory bear resemblance to indicator theory. An indicator is often defined as a measurable variable used to describe the status of a phenomenon (Øien, 2001), where the actual condition is unknown (Haugen et al., 2012). This perspective is often used in applied work within organisations seeking to monitor the status of risk influencing factors in the domain of interest.

This leads to our research problem:

Does the Safety Perception Index on Norwegian Vessel (SPIN-V) show preliminary satisfactory psychometric qualities, and how do the factors of SPIN-V relate to safety outcomes?

In our study, a preliminary validation of safety-related perceptions on Norwegian vessels, we conduct a series of analyses on the dataset, including dimension reduction techniques, reliability analyses and multivariate regression statistics.

Psychometric qualities are operationalized as (a) factor structure with theoretically and statistically sound factors, (b) factors with a satisfactory high internal consistency, (c) discriminant validity between factors, and lastly (d) convergent validity between relevant outcome measures.

2. The SPIN-V index

The NMA is the administrative and supervisory authority in matters related to safety of life, health, material values and the environment on vessels flying the Norwegian flag and foreign ships in Norwegian waters. The NMA has a strong focus on a risk-based inspection method, and have over the years developed a risk model, the aim of which is to help prioritize its joint efforts in order to reach peak efficiency.

The NMAs goal is to use its resources in such a way as to maximize the amount of safety and environmental protection achievable. In order to attain this goal, good quality data is of paramount importance. The NMA collects a large amount of data regarding accidents, inspections findings, vessel details, vessel activity, and more, in order to help prioritize its inspection and legislative efforts. This data is then used in the NMAs yearly risk assessment, which plays a key part in designing the NMAs causal accident models and make recommendations for further action.

Several of the NMAs accidents models for different specified incidents (both occupational

accidents and ship accidents) includes different organizational and human conditions contributing to a number of incidents. Upon gathering and processing all available data it was discovered that the NMA lacked high quality systematized data relating to these conditions. Upon reaching this conclusion the NMA initiated a project in order to gather and analyze the aforementioned data.

The NMAs accidents models are conceptual Bayesian Belief Network (BBN) models which consist of several nodes that have an impact on the probability of an accident (see e.g. Haugen et al., 2016). The items in the questionnaire was developed in order to get more information regarding the status of different nodes in the accident models. The selection of items was done in order to operationalize several of the nodes in the different models. This means that the questionnaire was not developed in order to operationalize one single construct such as safety culture, safety climate etc. As a consequence, the questionnaire consists of questions related to specified work practices (behavior), perceptions and attitudes related to work organization, safety management system (SMS), management (on the vessels and onshore), working environment, working conditions etc., as well as attitudes towards safety. The resulting survey data will be used in further developing the NMAs risk model, and in its yearly risk assessments. This will help the NMA to prioritize its inspection efforts, so that high-risk vessels will have a higher relative rate of inspections compared to low-risk vessels, as well as helping the NMA distinguish which elements to focus on for each individual inspection. In addition, the survey data will be used as a valuable source of information in various fact-finding projects.

The questionnaire was developed based on following steps:

i. Obtaining item sets from questionnaires applied in Norwegian ferry industry (Bye et al., 2015; Bye and Aalberg, 2020), shipping (Fagerholt et al., 2014), NORSCI (Tharaldsen et al., 2008) and safety climate questionnaire.
ii. Creating new items and modifications of existing items.
iii. Connection of items and selections as indicators of National Ship Risk Models (Haugen et al., 2016).
iv. Subject expert workshops with the NMA.
v. Pilot study in 2016, n = approx. 3000.
vi. Revising questionnaire based on pilot.
vii. Revising questionnaire based on input from the NMA, tripartite safety forum (SAFE) and trade union.

By basing our study on previous questionnaires, we hope to some extent accommodate that there are "many, many scales" of safety climate (Guldenmund, 2010).

3. Method

In the following section we provide descriptions of *procedure, sample* and *statistical analysis.*

3.1 *Procedure*

The questionnaire was distributed electronically by e-mail using Netigate in January 2019. Participants were able to respond both on cell phones, tablets or regular computers. The system allowed for reminders, and in total the respondents got two. There was a lottery incentive by participating, where lucky drawers got an iPad. Contact information and background data was collected from the NMA, trade unions, federations and a few shipping companies. All parties signed data processor agreements according to EU regulations. Around 1200 fishermen were contacted by cell phone and were able to respond to a smaller set of questions. These are not included in our analyses. The items utilized in our analyses, are Likert-scale items addressed to all respondents, which roughly equals half of the questionnaire. The items were measured on an ordinal scale range from 1 (completely disagree) to 5 (completely agree). Items included one option for "Do not know" which was removed in the analysis.

3.2 *Sample*

The sample of the present study is seafarers working on maritime vessels registered in Norwegian registers (N = 9344). A total of 1107 of these responded on telephone to a selected number of items and are not further used in the present study. Number of respondents subtracted missing items is 8095, with small variations across analyses. Norwegian males constitute almost all respondents.

3.3 *Statistical analyses*

3.3.1 Principal component analysis

Principal component analysis (PCA), referred to as "factor analysis", is a dimension reduction technique designed to look for common latent factors of measured items. In our study, we use varimax rotation, minimal number of iterates, and a combination of Kaiser's criterion <1 and scree

plot to determine number of factors. The PCA applied to questionnaires like this implicitly infer a balancing between theoretical perspectives and statistical perspectives in creation of factors. A combination of exploratory and confirmatory factor analysis was conducted, leading to the final 7-factor solution.

The results of the analysis were interpreted and validated according to the method recommended by Hair et al. (2014, pp. 95-114). Internal consistency (reliability) of the factors was examined by evaluating Cronbach's alpha and Dillon-Goldstein's rho (DG rho), in which the recommended level should be above 0.7. *Convergent validity* of the factors (i.e. if the factor adequately explains the variance in related variables) was assessed by use of partial least square structural equation modelling (PLS-SEM). PLS-SEM also helps to check the *discriminant validity* of the factors (i.e. whether the factors are sufficiently independent of each other) (Hair et al., 2014).

3.3.2 *Multiple regression*

Three multiple regression models were tested in order to study the factor's ability to predict seafarer's subjective safety perception and their involvement in accidents or near-misses.

Multiple linear regression model with safety perception (*All in all, how would you evaluate safety in your work situation?)* as outcome variable and all factors as predictors. The model's fit is examined by the adjusted *R* squared. Standardized beta coefficients show each factor's predictive power. Multiple logistic regression model with involvement in severe accidents (*Have you during the last 12 months been involved in one or several situations that developed into a severe accident?*) as outcome and all factors as predictors.

Multiple logistic regression model with involvement in near-misses (*Have you during the last 12 months been involved in one or several situations that could have developed into severe accidents, but didn't?)* as outcome and all factors as predictors. The logistic regression models' fit is studied by examining Nagelkerke pseudo R squared, whereas *odds ratio* along with the *p*-value represents the strength of the predictor's relationship to the outcome variable. Standardized *z*-scores show each factor's predictive power on the outcome variable (Mehmetoglu and Jakobsen, 2016, p. 170).

4. Results

4.1 *Factor analysis*

A combination of exploratory factor analysis (i.e. the PCA) with subsequent evaluation of the factors and a confirmatory factor analysis suggested 7 dimensions within 33 items (Table 1). 5 items were excluded from the confirmatory factor analysis as they failed to demonstrate minimum loadings and conceptual similarity within their factors. All dimensions except one (*Job demands*; 0.63) clearly showed satisfactory reliability with Cronbach's alpha above recommended level of 0.7. Internal consistency for this factor was still high when measured by Dillon-Goldstein's rho (see 3.2), which is by some considered to be superior to alpha (Liu et al., 2015). Items within each factor were qualitatively examined and found to be conceptually similar. The analysis has thus revealed a statistically and theoretically robust dimensionality within the instrument.

Table 1. Factors and factor loadings with respective reliability indicators (alpha).

Factor	Loadings	a	No. of items
(1) Safe work practices	0.51-0.73	0.85	9
(2) Management´s safety involvement	0.55-0.75	0.82	5
(3) Procedure compliance	0.63-0.73	0.81	5
(4) Competence and training	0.57-0.74	0.75	4
(5) Manning	0.57-0.81	0.76	4
(6) Job demands	0.65-0.72	0.63	3
(7) Safety system	0.70-0.80	0.73	3
Total			**33**

Note: Factor loadings < 0.4 are not reported.

The factor *Safe work practices* is related to the concept of mindful safety practices and consists of items like "I report when I see a dangerous situation". *Management's safety involvement* is concerned with the shipping company's responsiveness, involvement, feedback and commitment to safety on the vessels. An example item is "The company responds on conditions we report".

Procedure compliance concerns around compliance / violation of safety-related procedures, for example measured by the item "I sometimes violate the procedures to get the job done". It also includes one item that might be conceptualized as objective conflicts or work pressure, but it seems to load heavily on this factor thus being linked closely statistically. One important feature, to accommodate the criticism put forth by Bye and Aalberg (2020), the items concerned with procedures and safety system are contextualized with relevant examples for the

given position on board, e.g. "procedure for arrival at dock" for navigators.

The factor *Manning* consists of items regarding the judgement of satisfactory manning in terms of quantity (e.g. "we usually have higher manning than the stipulated safe manning for the vessel"). *Safety system* is in general related to the SMS on board the vessel. One example item is "I have easy access to procedures and instructions for my work". *Competence and training* relate to the perceived time and quality for training, e.g. "we have sufficient time for training of employees on board". *Job demands* concerns whether sleep, work conditions and quantitative work requirements are draining. An item in this factor is "I have too much to do".

4.2 *Descriptive statistics*

The factors' means computed as indices and standard deviations are presented in Table 2. The numbers reveal that all of the factors are rated fairly high indicating high safety in the sample.

Table 2. Means and standard deviations of the factors

Index	Mean	SD
(1) Safe work practices	4.34	0.64
(2) Management´s safety involvement	3.94	0.89
(3) Procedure compliance	3.64	1.03
(4) Competence and training	4.14	0.88
(5) Manning	3.60	1.12
(6) Job demands*	3.58	1.05
(7) Safety system	4.56	0.67

Note: * Reversed

4.3 *Correlation analysis*

All the factors' mean scores significantly correlated with each other at low to moderate level (see Table 3), suggesting that they all reflect different dimensions of a similar concept. *Management's safety involvement* shows a somewhat higher correlation with *Safe work practices*.

Table 3. Correlations between the factors

Factor	1	2	3	4	5	6
1	-					
2	0.60	-				
3	0.52	0.55	-			
4	0.51	0.47	0.39	-		
5	0.38	0.44	0.28	0.40	-	
6	0.38	0.40	0.40	0.42	0.36	-
7	0.43	0.40	0.29	0.40	0.27	0.23

Note: All correlations are significant at 0.001 level.

4.4 *Internal consistency, convergent and discriminant validity*

Coefficients for internal consistency, convergent and discriminant validity are reported in the Table 4. All DG rhos were exceeding the recommended threshold of 0.7, thus indicating a high *internal consistency* (Hair et al., 2014, p.102-103). Average variance extracted (AVE) was higher than the minimal level of 0.5 for all factors except one (*Safe work practices*), demonstrating acceptable *convergent validity* (Hair et al., 2014, p.103). None of the squared interfactor correlations was higher that AVE, suggesting high *discriminant validity* (Hair et al., 2014, p.105).

Table 4. Squared interfactor correlations with respective AVE and DG Rho

Factor n	1	2	3	4	5	6	7
1	1.00	0.37	0.23	0.28	0.17	0.15	0.22
2	0.37	1.00	0.31	0.24	0.22	0.17	0.19
3	0.32	0.31	1.00	0.20	0.13	0.20	0.12
4	0.28	0.24	0.20	1.00	0.18	0.16	0.19
5	0.17	0.22	0.13	0.18	1.00	0.14	0.10
6	0.15	0.17	0.20	0.16	0.14	1.00	0.06
7	0.22	0.19	0.12	0.19	0.10	0.06	1.00
AVE	**0.48**	**0.59**	**0.58**	**0.57**	**0.59**	**0.57**	**0.66**
DG Rho	**0.89**	**0.88**	**0.87**	**0.84**	**0.85**	**0.80**	**0.85**

4.5 *Multivariate analysis for convergent validity*

Examination of multiple linear and multiple logistic regression models revealed that factors predicted subjective safety perception and whether the respondents were involved in accidents and near-misses the last year.

4.5.1 *Subjective safety perception*

The linear regression model demonstrated a good fit with 46% variance by the predictors (Table 5). All the factors significantly predicted subjective safety perception. *Safe work practices* and *Management's safety involvement* violation were among the strongest single predictors. The overall model supported association between the identified dimensions of safety climate and subjective safety perception.

Table 5. Summary of multiple linear regression model with *Subjective safety perception* as outcome

Subjective safety perception	*Std. Beta*
(1) Safe work practices	0.27***
(2) Management´s safety involvement	0.17***
(3) Procedure compliance	0.14***
(4) Competence and training	0.10***
(5) Manning	0.11***

(6) Job demands[a]	0.08***
(7) Safety system	0.05***
*** $p > 0.001$	
F (7, 8087) = 1000.19, $R^2_{adj.}$ = 0.46, $p < 0.001$	

Note: [a] Reversed

4.5.2 *Involvement in accidents*

The logistic regression model showed that *Safe work practices*, *Procedure compliance* and *Competence and training* significantly increased the possibility for being involved in severe accidents the preceding year (Table 6). Among these, *Procedure compliance* was the strongest predictor. However, the model explained only a small portion of variance in the outcome variable (*pseudo* R^2 = 0.07). The model might be biased due to the low number of observations and skewed distribution (low % of respondents reporting accident involvement).

Table 6. Summary of multiple logistic regression model with *Involvement in accidents* as outcome

Involvement in accidents	*OR*	*z*
(1) Safe work practices	**0.76****	**-2.86**
(2) Management´s safety involvement	0.96	0.55
(3) Procedure compliance	**0.69*****	**-5.90**
(4) Competence and training	**0.87***	**-2.15**
(5) Manning	0.97	0.64
(6) Job demands[a]	0.94	-1.07
(7) Safety system	1.01	0.13
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$		
LL (7, 8087) = -1596.46, *pseudo* R^2 = 0.07, $p < 0.001$		

Note: [a] Reversed

4.5.3 *Involvement in near-misses*

The logistic regression model revealed that low *Safe work practices*, low P*rocedure compliance*, low *Competence and training,* high *Job demands* and high *Safety system* significantly increased the probability of respondents' reported involvement in near-misses. (Table 7).

Table 7. Summary of multiple logistic regression model with *Involvement in near-misses* as outcome

Involvement in near-misses	*OR*	*z*
(1) Safe work practices	**0.72*****	**-6.16**
(2) Management´s safety involvement	0.93	-1.78
(3) Procedure compliance	**0.64*****	**-13.54**
(4) Competence and training	**0.81*****	**-6.05**
(5) Manning	0.98	0.73
(6) Job demands[a]	**0.89*****	**-3.97**
(7) Safety system	**1.14****	**3.04**
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$		
LL (7, 8087) = -4377.48, *pseudo* R^2 = 0.16, $p < 0.001$		

Note: [a] Reversed

5. Discussion

As pointed out initially, our purpose of the present study was to conduct a preliminary validation of the SPIN-V index by focusing on a) factor structure, b) internal consistency, c) discriminant validity and d) convergent validity.

In general, the factor structure provided logical solutions with robust measures of internal consistency. The factor *Job demands* had a somewhat low Cronbach's alpha, but acceptable rho.

As for discriminant and convergent validity, several notes can be made. First, we conducted correlations and squared inter-correlations. Correlations between factors were low to moderate, which is neither too high (collinearity) nor too low (no convergence). One could expect associations between concepts due to their theoretical resemblance. The associations the factors had with *Safety perception* might seem a bit high, indicating that there is one common factor that explains a large portion of safety perception. To conclude, there is one common "safety perception" factor that explains a lot of variation, but there are increments of variation on other factors as well. The convergent validity was satisfactory except for *Safe work practices*. It can therefore be hypothesized that safe work practices bear strong theoretical or statistical resemblance to one or more of the factors, and quite little to some of the others. As for the discriminant validity, average variance explained was not higher than intersquared correlations. It can thus be concluded that the factors show high internal homogeneity combined with high external heterogeneity.

Furthermore, we conducted three multivariate analyses for convergent validity; safety perceptions, involvement in accidents, involvement in near-misses. All models were significant, indicating an overall convergence with relevant outcome measures.

Safety perceptions was strongly predicted (46%) especially by *Safe work practices*, which again might indicate some resemblance of theoretical and/or test metric nature. The strong prediction might indicate that the notion of subjective safety perception indicates a perceived control over risk – by having strong tendencies of safe work practice. It is also interesting that the subjective safety perception is influenced/associated by perceptions of management's involvement, which indicate that the framework conditions indirectly or directly enforced by the shipping company plays a significant role in the overall notion of safety in the seafarers' work.

Prediction of involvement in accidents were significant, but low (7 %). High *Safe work practice*, high *Procedure compliance*, and high *Competence and training* might be considered as protective phenomena for involvement in accidents. There is, however, high uncertainty in this line of thought due to the cross-sectional nature of our study.

As for involvement in near-misses, the model was a stronger predictor (16 %) than for accidents. The same predictors were significant, in addition to *Job demands* and *Safety system*. It is interesting that high *Job demands* seem to influence involvement of near-misses but not accidents. It might be due to low statistical power or due to amount of accidents in the study. The relationship might also be related to that near-misses are more closely related to sharp end work pressure, whereas accidents are more of a complex nature and that barriers might hinder work pressure solely leading to accidents. A surprising result was the relationship between higher satisfaction with safety system on board, and involvement in near-misses. An explanation might be that a low quality or non-existing SMS, accompanied with an absence of reporting culture, might lead to that the respondent do not consider actual near-misses as near-misses, and therefore do not report them.

6. Concluding remarks

Our preliminary assessment of the construct validity of the SPIN-V index shows promising results in using the perceptions of seafarers as indicators of safety-related factors in the maritime industry. The factor *Job demands* might need further development to increase reliability. Further study should assess different aspects of the validity of the questionnaire, such as predictive validity and using multitrait-multimethod procedures.

A strength of our study is the high number of responses compared to previous studies of maritime personnel's safety perceptions, being one of the largest studies of its kind to date. Further application of the material in this data should address differences between types of vessels, since the maritime industry is heterogenous and with radically different operational patterns, and include more factors and control variables in multivariate analyses. As a questionnaire targeting safety climate and organisational culture is only directed to "espoused values", qualitative methods are recommended for further endeavour into the relationships identified. Further studies should also explore other parts of the questionnaire data, including reporting of navigation-related aspects, technical conditions, framework conditions like shift arrangements, and cross-cultural crew communication.

References

Arslan, V., Kurt, R. E., Turan, O., and De Wolff, L. (2016). Safety culture assessment and implementation framework to enhance maritime safety. Transportation research procedia, 14, 3895-3904.

Bergheim, K., Nielsen, M. B., Mearns, K., and Eid, J. (2015). The relationship between psychological capital, job satisfaction, and safety perceptions in the maritime industry. Safety science, 74, 27-36.

Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281. doi: 10.1037/h0040957

Darbra, R. M., Crawford, J. F. E., Haley, C. W., and Morrison, R. J. (2007). Safety culture and hazard risk perception of Australian and New Zealand maritime pilots. Marine Policy, 31(6), 736-745.

Ek, Å., and Akselsson, R. (2005). Safety culture on board six Swedish passenger ships. Maritime Policy and Management, 32(2), 159-176.

Ek, Å., Runefors, M., and Borell, J. (2014). Relationships between safety culture aspects–A work process to enable interpretation. Marine Policy, 44, 179-186.

Ek, Å., Olsson, U. M., and Akselsson, K. R. (2000, July). Safety culture onboard ships. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, No. 27, pp. 320-322). Sage CA: Los Angeles, CA: SAGE Publications.

Engvik, H. (1999). Testvaliditet. *Tidsskrift for Norsk Psykologforening, 36*, 718-722.

Guldenmund, F. W. (2010). Understanding and exploring safety culture. Uitgeverij BOXPress.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Haugen, S., Almklov, P. G., Nilsen, M., and Bye, R. J. (2016). Norwegian national ship risk model. In Proceedings of the 3rd International Conference on Maritime Technology and Engineering-MARTECH 2016. CRC Press.

Håvold, J. I. (2010a). Safety culture and safety management aboard tankers. Reliability Engineering and System Safety, 95(5), 511-519.

Håvold, J. I. (2005). Safety-culture in a Norwegian shipping company. Journal of safety research, 36(5), 441-458.

Håvold, J. I., and Nesset, E. (2009). From safety culture to safety orientation: validation and simplification of a safety orientation scale using a sample of seafarers working for Norwegian ship owners. Safety Science, 47(3), 305-326.

Håvold, J. I. (2010b). Safety culture aboard fishing vessels. Safety science, 48(8), 1054-1061.

Haavold, J. I., Nesset, E. R. I. K., and Strand, O. (2011). Safety attitudes and safety ambivalence among officers from the Philippines and Norway. Safety Science Monitor, 15(1), 1-15.

Håvold, J. I. (2015). Stress on the bridge of offshore vessels: examples from the North Sea. Safety science, 71, 160-166.

Håvold, J. I. (2007). National cultures and safety orientation: A study of seafarers working for Norwegian shipping companies. Work and Stress, 21(2), 173-195.

Liu, H., Bahron, A., and Bagul, A. H. B. P. (2015, December). Justifying scale type for a latent variable: Formative or reflective?. In AIP Conference Proceedings (Vol. 1691, No. 1, p. 050009). AIP Publishing LLC

Lu, C. S., and Tsai, C. L. (2008). The effects of safety climate on vessel accidents in the container shipping context. Accident Analysis and Prevention, 40(2), 594-601.

Mehmetoglu, M., and Jakobsen, T. G. (2016). *Applied statistics using Stata: a guide for the social sciences*. Sage.

Murphy, K. R. and Davidshofer C. O. (2005). *Psychological testing: Principles and applications.* (6. Utg.). New Jersey: Person Education Inc.

Nævestad, T. O. (2017). Safety culture, working conditions and personal injuries in Norwegian maritime transport. Marine Policy, 84, 251-262.

Oltedal, H., and Wadsworth, E. (2010). Risk perception in the Norwegian shipping industry and identification of influencing factors. Marit. Pol. Mgmt., 37(6), 601-623.

Shang, K. C., and Lu, C. S. (2009). Effects of safety climate on perceptions of safety performance in container terminal operations. Transport reviews, 29(1), 1-19.

Tharaldsen, J. E., Olsen, E., and Rundmo, T. (2008). A longitudinal study of safety climate on the Norwegian continental shelf. *Safety Science*, *46*(3), 427-439.