



PSYCHOMETRIC TESTS:

MEASURING THE PSYCHOLOGY BEHIND
GOOD PERFORMANCE ON BOARD



WHITE PAPER

05/2019



Psychometric tests:

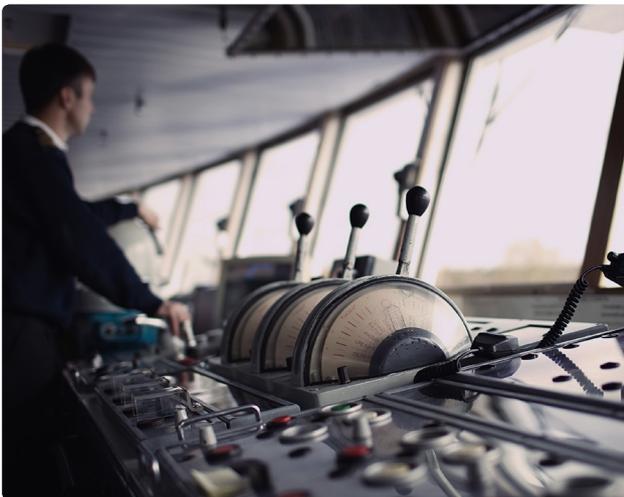
Measuring the psychology behind good performance on board

The use of psychometric tests has been on the rise during the last decades and the maritime industry is quickly catching up. With the increased recognition of the importance of psychological attributes such as cognitive abilities and “soft” skills for the optimal performance on board, the quest for measuring the psychology behind good performance has set off.

This white paper looks at the important psychological attributes in detail, the characteristics of the different types of tests that measure them and how to choose the ones that are best suited for assessing your ships’ crew.

How does peoples' psychology affect performance on board?

People are the pumping heart of the maritime industry. Their characteristics and behaviour are well recognised to be crucial for ensuring maritime safety, security and marine environmental protection. Both the Standards of Training and Watchkeeping for Seafarers (STCW) and the International Safety Management (ISM) code were developed to introduce standards and procedures, recognising the important role of crews on board.



However, for many years regulations focused mainly on the technical knowledge and skills of seafarers.

The so-called Manila Amendments to the STCW included in 2010 for the first time non-technical skills such as leadership and teamwork. It introduced also a new provision related to obtaining and maintaining situation awareness, which also is related to the psychological domain.

In a similar vein, many Maritime Education and Training (MET) organisations started offering bridge/engine resource management courses, in an attempt to increase the understanding of how human psychology affects performance at sea.

Psychological attributes affect the performance indirectly, by influencing how they perform their technical tasks and directly, by complementing technical skills. An engineer is expected to have both the skills to perform certain maintenance tasks on board, but also the persistence to identify the cause of the problem and keep looking for solutions until the task is solved or help is received from ashore.

Furthermore, the engineer must be able to be assertive in his communication with the rest of the crew to make sure that everyone understands the problem and its implications for the related operations.

Human-error and accident investigations

The importance of psychological skills becomes particularly visible during accident investigations. Several studies indicated that human error is among the most common reasons leading to accidents in the maritime industry. It was found to be the main reason for collisions, fires and other types of accidents that cause losses of cargos or human life.

Different sources estimate that the human factor is the main reason for up to 80–85% of maritime accidents. The Standard P&I Club estimates that over a recent ten-year period, insurance claims cost the P&I industry 15 billion dollars. Over 65% of this pay-out – 10 billion dollars – was for incidents in which humans played the dominant part¹.

“Human error is among the most common reasons leading to accidents in the maritime industry”

Of course, not all human errors are related to psychological factors. There have been many accidents clearly attributable to lack of proper training or education. However, there are also other occasions where lack of technical knowledge was ruled out, focusing instead on psychological attributes.

For example, after Costa Concordia capsized, the Italian authorities concluded that the root cause of the accident was the failure of decision-making and coordination that caused the contact with the rocks and the poor emergency management². The incident resulted in the death of 32 persons and the injury of 157 others, as well as the loss of the ship and significant environmental damage, summing up to 2 billion dollars.

1. Maritime and Coastguard Agency (2010). The Human Element: a guide to human behaviour in the shipping industry.
2. Italian Ministry of Infrastructures and Transports (2013). Costa Concordia: report on the safety technical investigation.

Which are the psychological attributes that influence performance on board?



The issue of how the psychology of individuals affects their performance is as old as humankind. 3000 years ago, in China, the Emperor not only assessed officers on the basis of the Six Skills (arithmetic, archery, horsemanship, music, writing and skills in the performance of ceremonies), but also the Six Conducts (filial piety, friendship, harmony, love, responsibility and compassion) and the Six Virtues (insight, kindness, judgement, courage, loyalty and concord)³.

Today, the framework did not change dramatically, though we have more elaborate theories about which psychological attributes affect performance and in what way.

These theories usually fall within the following three categories:

Soft skills

There are many ways to refer to the competencies that fall outside of one's "hard", technical expertise: soft skills, social and emotional skills, interpersonal skills, self-management, emotional intelligence (EI), life skills. While the different terms accentuate different aspects; they refer to an overlapping set of knowledge, skills and attitudes. This proliferation of applicable terms reflects the varying set of non-technical skills required by each sector.

Specific to the maritime industry, leadership and teamwork have been identified by STCW as important competencies for officers on board, however, those definitions are too general to inform concrete actions. In fact, due to the difference in the definitions and theories behind those skills, it was found that MET organisations offer training programmes that are not consistent with each other and send contradictory messages to officers.⁴

Another source of information as to which soft skills are relevant to the maritime industry, are the claims analyses published by the P&I clubs. For instance, the Swedish Club⁵ identified the following deficiencies of soft skills that caused accidents:

3. Rust J. and Golombok, S. (2009) Modern Psychometrics, Third Edition: The Science of Psychological Assessment.
4. Mori Y. (2014) - An Analysis of leadership education and training in maritime education and training institutions.
5. The Swedish Club (2016). P&I Claims Analysis.

- Lack of belief in safety and over-confidence in one's ability.
- Lack of communication between crew members.
- Poor communication between crew and office staff.
- Not acknowledging cultural differences between nationalities, company and profession.
- Not being assertive when mistakes have been made by others.

One limitation of the claims analyses is that they focus only on situations where things went wrong and do not take into account the skills that seafarers possess in order to make things right every day. After all, accidents are rare and given the number of voyages, good performance is the norm in the industry rather than an exception. This underscores that importance of understanding not only why humans fail, but also why they succeed.

Cognitive ability

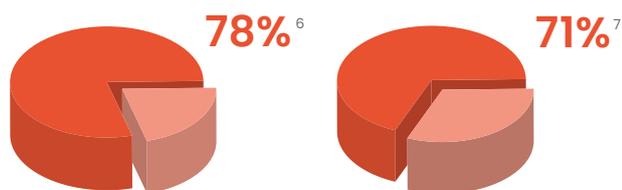
The cognitive ability or cognitive skills refer to the ability of people to acquire, perceive, process, store and project information. The specific cognitive abilities, such as memorisation, spatial perception, visualisation, are also called aptitudes. The cognitive ability affects how much and how fast people learn new information. The amount of job-specific knowledge to be acquired is significant for all job levels, but especially for more advanced positions. This is the reason why cognitive ability is essential for success in high-complexity jobs.

Furthermore, cognitive ability affects the performance directly. A lot of cognitive abilities, e.g. problem sensitivity for identifying existing or potential issues, are exercised and needed on a daily basis in many jobs.

While STCW explicitly states that "cognitive skills [...], to varying degrees, underpin all levels of competence", the regulation does not go into details as to what are the specific cognitive skills that are relevant to the performance of seafarers. The International Regulations for Preventing Collisions at Sea also emphasise the importance of cognitive skills, though not naming them as such, mentioning „making a full appraisal of the situation“ or „not making assumptions based on scanty information“.

The concept of situation awareness is related strongly to the concept of cognitive ability, as it is defined as 'being aware of what is happening around you and understanding what that information means to you now and in the future'⁵, which refers to the perception, processing and projection of information, respectively.

Accidents due to human errors



Loss of situation awareness has been implicated in a number of accidents at sea as indicated in the following case studies:

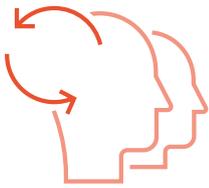
An analysis of 23 accident reports for collisions found that in 18 out of 23 cases, loss of situation awareness precedes the accident⁶.

5. Endsley, Mr. R. (2012). Designing for Situation Awareness
6. Sandhåland, S. et al. (2015). Situation awareness in bridge operations – A study of collisions between attendant vessels and offshore facilities in the North Sea.

An analysis of 177 maritime accident reports (public domain) originating from eight different countries found that 71% of human errors were situation awareness related problems⁷.

Among the other factors that influence the situational awareness, it has a strong cognitive component and soft skill aspect related to the communication within the team.

Personality



Personality refers to individual differences in characteristic patterns of thinking, feeling and behaving. Personality is considered an enduring characteristic of individuals that varies little throughout their lifespan, even when faced with major changes in life circumstances.

At the same time, people do not show the same personality trait to the same degree through their entire life. For example, a talkative, sociable third officer will very likely turn into a lively, well-connected chief mate, and while it is unlikely that he/she will retain the same level of sociability at the age of 80, he/she may still be above average for his/her peers.

It seems that personality is a good predictor of behaviour in situations where people have freedom of choice. However, personality differences do not explain well differences in performance, in situations when people have to comply with strict regulations or follow specific demands⁸.

Research on the role of personality in career success, suggested that people with certain personality characteristics function best in certain environments. If they work in their preferred environment, this is associated with better performance, organisational citizenship behaviour and persistence⁹.

There is a lack of systematic evidence that relates performance of seafarers on board to certain personality characteristics.

7. Grech, Horberry & Smith (2002). Human Error in Maritime Operations: Analyses of Accident Reports using the Leximancer Tool.
8. Specht, J. (2017). Personality Development Across the Lifespan.
9. Nye et al. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research.

How to measure psychological attributes?

Recognising that individual differences exist is an important step towards ensuring that people are recruited for positions where they can utilise their strengths. However, this cannot happen without using means to measure these individual differences in a valid and reliable way.

Self-report tests



Self-report tests or inventories are the most prevalent forms of tests used to assess soft skills or personality. This type of test is presented in a paper-and-pencil format or may even be administered online.

A typical self-report inventory presents a number of questions or statements that may or may not describe certain qualities of the test subject. For example, test takers may be asked to what extent they agree with statements like “I find it easy to organise the schedule of my team” in order to assess their coordination skills.

Self-report tests have the advantage of being:

- an inexpensive option to assess the psychological qualities of a large number of participants, especially if administered online and scored automatically
- fair, because the subjects of the test report about themselves, so to speak, which means that the information comes straight from the horse’s mouth
- consistent, as people are generally more motivated to give information about themselves, than about others.

Performance tests

Performance tests are mainly used to assess cognitive ability. They involve a standardised administration of tasks that measure some aspects of cognitive functioning. The tasks within the tests are usually timed and have right and wrong answers. The speed of the responses and their accuracy is taken as an indication of the degree to which the person possesses the respective cognitive ability.

For example, a typical assessment of memorisation is presenting several objects for a limited time, asking the test taker to recall the objects he/she saw. Nowadays many of the available tests are computer-based or browser-based. The advantages of performance tests are that:

- they provide an objective measure of one's ability based on their speed and accuracy and not on someone's perception of their speed and accuracy;
- they are mostly unaffected by language, which makes such tests suitable for the multi-language population;
- their large-scale administration is cost-efficient, especially if administered online and scored automatically.

The disadvantage is that the tasks assess latent structures, such as working memory or attentional span, so the test takers might not perceive it as relevant to their day-to-day tasks.

Also, speed and accuracy are affected significantly by the current motivation of the test taker, if the environment allows for good concentration and if he/she is well rested.

Behaviour observations/interviews

Behaviour observation is a widely used method for the assessment of observable psychological attributes. It involves watching and recording the behaviour of a person under certain circumstances.

Behaviour can be rated using dimensions such as quality (low-high) or a relative value (below average – average – above average, see Fig.1). For complex behaviour, numerous positive and negative indicators can be provided, to guide the evaluation of the observed behaviour.

For example, a positive indicator can be: "has a logical scheme for deciding what is important" and a negative indicator might be: "takes decisions by himself/herself".

The advantages of the behaviour observation/interview as a method are:

- it is an ecologically valid way to assess observable psychological characteristics such as soft skills, since the behaviour of the person is directly observed during the performance of his everyday tasks or is reported by the person himself;
- generally, there is a good acceptance among employees for the results of such assessments.

The major disadvantages are that:

- the method is very resource-consuming;
- the people who are going to do the behaviour observation or the interview should be well trained to notice the assessed behaviour;
- despite proper training, they are still biased by their own beliefs and limitations.
- this can be especially problematic when you have to assess diverse workforce like seafarers.

How are different kinds of tests scored?

Objective versus subjective scoring

A test is considered objective when the answers are recorded and directly converted into the final scoring.

Cognitive tests are an example of such tests where the response time and the number of accurate responses is recorded and either presented as they are or converted into scores based on pre-existing norms.

A subjectively scored test is the one that requires the examiner to convert the response of the test taker into a score. These are typical behaviour observation methods. Additionally, there are interviews, where the examiner interprets the behaviour in order to provide scoring.

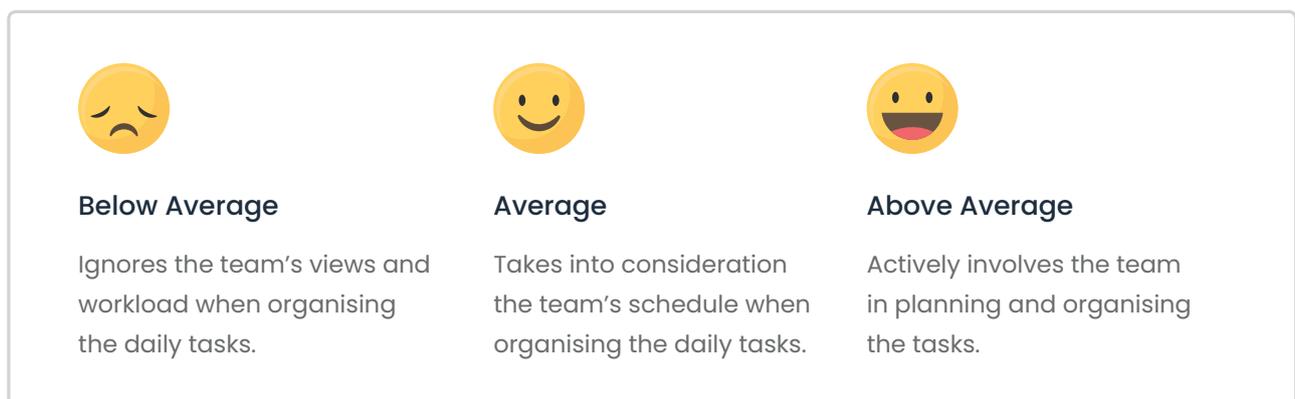


Fig. 1

Norm-referenced vs criterion referenced scores

Most psychometric tests are norm-referenced, which means that they provide info on how the test takers' rankings are compared to each other. For example, a seafarer may have scored in the 97th percentile, which means that he did better than 97% of the same norm group.

There is no attempt to interpret the scores in terms of what the test taker knows precisely, but it gives information whether his performance is typical of low-, middle- and high-performance test takers in the norm group. In order to simplify the scoring, sometimes instead of percentiles, STANINE (Standard nine) scores from 1 to 9 are used. One stanine score refers to a particular range of percentiles. (see Fig. 2).

Because the scoring is always relative to the norm group, it is crucial to make sure this group has the characteristics of a group you are interested in assessing. For example, measuring the coordination skills of a captain with a reference to a norm group consisting of ratings, will yield a much higher score than if he/she is measured against other captains.

Criterion-based scoring, on the other hand, indicates if the test taker's performance meets a certain standard, desirable performance, benchmark or criterion.

The set of behaviours that the test taker must demonstrate is specified by a committee of experts that conduct opinion surveys and observations to set the standard. These usually are "pass or fail tests" and, they are rarely used for assessing psychological characteristics.

How is a psychological test being used?

Selection



A psychometric assessment can be very beneficial in the selection process since it can improve the accuracy of selection. Through a psychometric assessment, the candidates are "objectively" assessed and thus companies gain additional information in order to make good hiring decisions.

Cognitive ability, soft skills and personality tests can all be used for the selection process. Administering a combination of tests usually gives more reliable results than conduction of a single test.

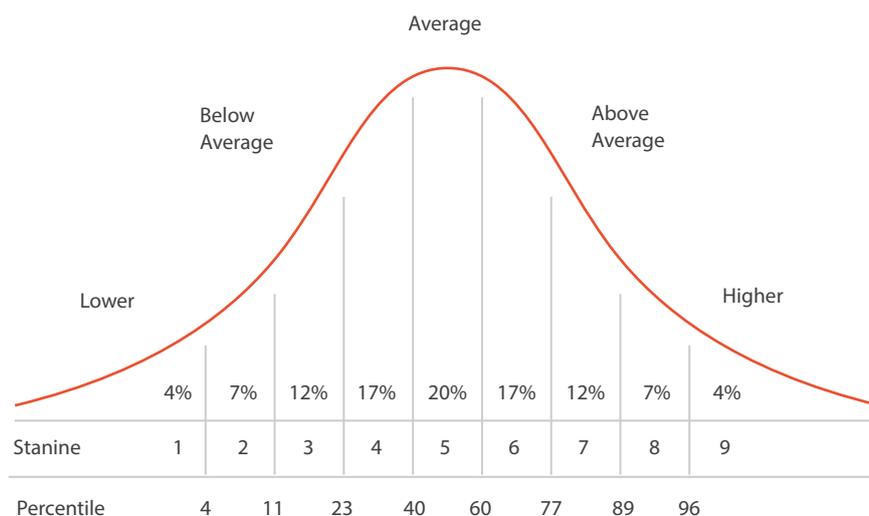


Fig. 2.

Promotion



Companies may also use psychometric assessments in order to identify the employees who possess higher level abilities, so that these employees can be promoted and take over more responsibility.

For example, the result of a soft skills test can indicate that a particular officer is ready to be promoted from operational to managerial level.

Training and development



Companies are increasingly using psychometric assessments for training and developmental purposes. Psychometric tests can help identify specific areas that require training in order to design and implement well-targeted training programmes, rather than assuming that everyone needs the same training.

For example, through the use of soft skills assessments, a company may find out that 30% of the participants would benefit from training in “self-control” skills and assign a proper training specifically for them and not generic for everyone.

Personality tests can be used to develop specialised training for people with different personality profiles. Standardised interventions show different effectiveness across individuals. Customised programmes are one way to optimise the impact of the interventions, when taking into consideration the personality factor.

Personality traits are stable across the lifetime and for this reason they, unlike soft skills, are not an appropriate target for modification.

Cognitive tests are usually not used for training and development because cognitive characteristics are rather stable across the lifespan¹⁰.

10. Ronnlund, Sundstrom, & Nilsson (2015). Interindividual differences in general cognitive ability from age 18 to age 65 years are extremely stable and strongly associated with working memory capacity.

How to choose a high quality test?

“The main criteria for a high-quality test are: reliability, validity and fairness of the test.”

There are thousands of tests out there that claim to measure psychological characteristics. What is the difference between a legitimate psychometric assessment and a questionnaire posted in a magazine? The main criteria for a high-quality test are: reliability, validity and fairness of the test¹¹.

Reliability

Reliability is a valuable property that every assessment should demonstrate in order to be useful and effective. A reliable instrument will give accurate and consistent scores across repeated assessments, given that the psychological characteristics have remained stable. For example, the results of cognitive ability or personality test should remain stable over several years.

Cognitive abilities, however, get affected by chronic stress, presence of psychological disorders and other pathologies¹². Similarly, the results of a soft skill assessment should be stable over 1–2 years, given that meanwhile, the test taker did not undergo any intervention.

Another type of reliability is the so-called internal consistency. The individual's responses to various parts of the same test are examined to check if they are in agreement and measure the same skill/area consistently and accurately. If ten statements in a test measure “coordination” skills, it is expected that the test taker with high coordination skills will rate all or most of them higher.

Validity

Another important property that every test should have is validity. This criterion concerns the extent to which the test measures what it should measure. A test can be reliable, but not valid. For example, a test can measure height correctly every time, but it is still an invalid way to assess someone's memory.

11. Ronnlund, Sundstrom, & Nilsson (2015). Interindividual differences in general cognitive ability from age 18 to age 65 years are extremely stable and strongly associated with working memory capacity.
12. Marin M. F., et al. (2011). Chronic stress, cognitive functioning and mental health.

There are several types of validity:

Face validity:

It refers to how a test appears to be measuring the skill/attribute that it is supposed to measure. For example, a test that measures memory by asking people to remember strings of letters has higher face validity than a test that assesses memory by asking people to run fast on a treadmill.

Construct validity:

This refers to the extent to which a test measures the psychological construct or characteristics that it is intended to measure. If a cognitive ability test is recommended for use with engine officers, it is important to be able to demonstrate that the cognitive abilities are the ones that are relevant to the performance of the engineers.

Life on board is different than the work environment at the office, and the same skills can be manifested in different ways. For example, the coordination skills of a captain, are associated with different behavioural indicators than the coordination skills of a project manager. For this reason, the test content must not only be adjusted to make sure the desired psychological characteristics are measured, but also to the profession and the specific conditions of this profession.

Criterion-related validity:

It examines the correlation or any other statistical relationship between the individual's performance on a test and the actual job performance of the same individual. For example, individuals who score high on a soft skills test tend to receive higher job appraisals for their soft skills than those who score low on the test.

Predictive validity, on the other hand, indicates to what extent the individual's test performance predicts his/her performance in the future. Both predictive and criterion validity are challenging to demonstrate, because they require test scores to be compared to a valid objective and a reliable measurement of work performance, which is often lacking.

Test fairness

An essential factor that should be taken into consideration when selecting a test is its fairness. Proper psychometric assessments should not be discriminating, instead they should be fair to everyone irrespectively of gender, culture, age, race and other factors.

A fair application of a test also means that the employee is assessed on skills and abilities that are relevant and important for the performance on the job.

Conclusion

Psychometric assessments, when applied appropriately, can serve as an essential component of the crewing process and increase the safety and efficiency of the operations on board. However, the choice of an instrument should not be taken lightly as there are a lot of factors to be considered: the available resources, the desired application, the characteristics of the test itself and the scoring.

Also, it is important to remember that the combination of different methods should be used in order to make sure the assessment of the individual is valid, reliable and fair.

A note of caution

All methods used to assess psychological characteristics have their advantages and disadvantages. Inevitably some errors occur with each measurement. For this reason, more than one method should be used in order to obtain more accurate results.

This refers to the principle of aggregation that states that the sum of a set of multiple measurements is a more stable and reliable estimator of a personality trait, skill or ability than any single measurement¹³.

*The discussion of tests with clinical application is beyond the scope of this paper.

13. McDonald (2008). Measuring Personality Constructs: The Advantages and Disadvantages of Self-Reports, Informant Reports and Behavioural Assessments.

About SafeMetrix

The SafeMetrix platform powered by Safebridge was developed in cooperation with seafarers and some of the leading experts in the industry. Its core purpose is to offer maritime-specific crew assessments, which allow you to measure specific skill sets of seafarers and provide the users with otherwise hard-to-obtain data for better-informed decisions.

The integration of our platform into the crewing process provides substantial benefits to both companies and seafarers operating in the maritime industry.

This applies to the following areas:

Selection

Enhanced validity of interviews via preliminary screening

Retention

Improved position adaptability and decreased turnover

Promotion

Enhanced timely and transparent career development

Training

Optimised training initiatives with targeted actions

Assessment Portfolio

The SafeMetrix portfolio is currently focused on delivering maritime-specific psychometric assessments for the crews' non-technical skills. These skills are increasingly important for both the wellbeing of the seafarer and the overall safety of the operations at sea.

MET-3S: Soft Skills for Seafarers

MET-3S is a psychometric assessment, specifically developed for the maritime industry. The product offers insights into the "soft skills" of officers on board across 13 skills grouped in 3 clusters:

- Coping under pressure
- Interpersonal skills
- Self-management

Compared to other psychometric assessments, MET-3S was developed specifically for assessing the competence of crews on board. The tested skills are selected after a thorough job analysis of the different ranks on board to match the requirement of the position.

MET-CSS: Cognitive Skills for Seafarers

MET-CSS is a psychometric assessment, specifically developed for the maritime industry. The product offers insights into the “cognitive skills” of officers on board:

- Memorisation
- Problem sensitivity
- Spatial orientation
- Information ordering

Compared to other psychometric assessments, MET-CSS was developed specifically for assessing the competence of crews on board. The tested skills are selected after a thorough job analysis of the different ranks on board to match the requirement of the position.

This psychometric assessment constitutes a performance test based on the speed and accuracy of the responses of the test taker.

SafeMetrix Key Features:

- ✓ Accessible Online
- ✓ User-Friendly Interface
- ✓ Intuitive Dashboards
- ✓ Immediate Reporting
- ✓ 24/7 Customer Support

Visit www.safemetrix.com for more details on SafeMetrix.



WHITE PAPER

05/2019