The Potential of Maritime-Scientific Numeracy Tasks for Developing Key Competencies of Seafarers

Agus Dwi Santoso[1*], Achmad Dhany Fachrudin[2]

[1,2] Politeknik Pelayaran Surabaya, Surabaya, Indonesia

Email: agusbp2ipsby@gmail.com

The digital transformation of the maritime sector creates an urgent need for seafarers to master advanced cognitive skills (reasoning, problem solving, critical thinking, data processing, and STEM application). This study examines whether numeracy tasks grounded in maritime scientific contexts can elicit those key competencies. Motivated by gaps in current MET curricula and the close link between numeracy and seafaring decision-making, we developed six tasks across four subcontexts (navigation, marine engineering, oceanography, maritime economics). Two domain experts rated each item using a competency-based rubric; data were analyzed with descriptive statistics (means, SD) and inter-rater reliability (ICC), complemented by general student feedback. Overall, findings indicate that the developed numeracy tasks successfully elicited key seafarer competencies, as evidenced primarily by expert evaluations and supported by general student perspectives. Specifically, STEM Knowledge & Application and Information & Data Processing were consistently and strongly elicited, while Problem & Conflict Solving showed moderate potential. Reasoning & Decision Making and Critical Thinking produced inconsistent expert agreement (low ICC), indicating rubric ambiguity and assessor variability. We recommend integrating maritime-scientific numeracy tasks, particularly data- and STEM-rich items, into MET mathematics curricula, along with rater-calibration sessions and expansion of rater and sample pools to improve reliability and validity. Overall, maritime-context numeracy tasks are promising dual-purpose instruments for both assessment and learning in competency-based maritime education.

Keyword: *Maritime education and training (MET); Numeracy Tasks; Key Competencies of Seafarers;*

1. INTRODUCTION

The maritime sector is experiencing accelerated digital transformation, driven by automation, intelligent navigation systems, and data-based vessel management, which fundamentally reshapes the competency requirements for seafarers. Modern maritime operations demand complex cognitive abilities such as reasoning and decision making, problem and conflict solving, critical thinking, information and data processing, and STEM knowledge and application (Ceylani et al., 2022). These skills are inherently linked to mathematical literacy (numeracy), which enables seafarers to interpret

M. A. Rohman et al. (eds.), *Proceedings of the International Conference of Inland Water and Ferries Transport Polytechnic of Palembang on Law, Economic and Management (IWPOSPA-LEM 2025)*, Advances in Economics, Business and Management Research 396,
https://doi.org/10.2991/978-94-6239-729-3_18

navigational data, calculate optimal routes, manage energy efficiency, and make operational decisions based on quantitative evidence (Stanivuk et al., 2017).

However, current international frameworks, particularly the Standards of Training, Certification and Watchkeeping (STCW), have not yet fully adapted to these evolving demands. Ghosh et al. (2024) emphasized that the STCW framework "may have to reinvent itself" to include cognitive and digital skills relevant for the seafaring workforce of the future. Similarly, Belabyad et al. (2025) underlined the growing gap between existing training standards and the competency requirements emerging from autonomous and smart-ship operations. This mismatch underscores an urgent need to develop instructional and assessment tools that explicitly target higher-order cognitive and digital literacies for maritime professionals.

Recent studies demonstrate that integrating authentic disciplinary contexts into mathematical tasks strengthens transfer, engagement, and conceptual understanding. Contextualized numeracy, particularly in maritime settings, has been shown to enhance analytical thinking and the application of quantitative reasoning to real-world marine operations (Mycroft & Sriraman, 2020; Zaura et al., 2025). Research by Fachrudin et al. (2025) introduced a conceptual framework for maritime-scientific numeracy encompassing navigation, marine engineering, oceanography, and maritime economics, revealing strong potential for aligning mathematics learning with seafaring practice. Nonetheless, few studies have empirically validated whether such tasks genuinely elicit the key cognitive competencies defined for future seafarers.

This study addresses this gap by developing and validating maritime-scientific numeracy tasks designed to elicit five target competencies: reasoning and decision making, problem and conflict solving, critical thinking, information and data processing, and STEM knowledge and application. The innovation of this research lies in three aspects: (a) the creation of tasks grounded in four authentic maritime subcontexts, (b) systematic expert mapping using an analytical rubric derived from Niss & Hojgaard (2011) mathematical competency framework, and (c) methodological triangulation between expert judgments, inter-rater reliability analysis (ICC), and general student perspectives.

Using a descriptive analytical approach, six tasks were evaluated by two experts in mathematics education and maritime studies. Each task was scored against fifteen competency indicators. Descriptive statistics and Intraclass Correlation Coefficient (ICC) analysis (Koo & Li, 2016) were applied to examine inter-rater reliability, complemented by general observations of student engagement to provide practical context. The research aims were to (1) determine which key seafarer competencies are most effectively elicited by maritime numeracy tasks, (2) complement and contextualize expert assessments through general student perspectives, and (3) provide empirical recommendations for aligning maritime numeracy assessment with future competency needs.

2. METHODOLOGY

The study used a descriptive analytical approach to examine whether a set of six maritime-scientific numeracy tasks elicits five targeted seafarer competencies. The tasks, constructed to cover four maritime subcontexts (navigation, marine engineering/stability, oceanography, and maritime economics), were developed from

the maritime-numeracy framework described in recent literature (Fachrudin et al., 2025). Two subject-matter experts (one in mathematics education and one in maritime education) independently rated every task using a 15-item rubric that maps three observable indicators to each competency domain (Reasoning & Decision Making; Problem & Conflict Solving; Critical Thinking; Information & Data Processing; STEM Knowledge & Application). Expert ratings used a 1–5 ordinal scale and were aggregated to compute competency means and item summaries per task (Fachrudin et al., 2019; Niss & Hojgaard, 2011).

To complement the expert evaluations, general feedback and observational insights were gathered from a pilot group of undergraduate students who interacted with the tasks. Rather than employing a formal deductive coding scheme, this phase focused on capturing general student perspectives regarding task clarity, engagement, and the practical application of the targeted skills (e.g., justifying steps, applying domain formulas, or interpreting data). This informal observational process sought to provide practical context, exploring whether the students' general approaches aligned with the competencies predicted by the expert raters.

Quantitative analysis comprised descriptive statistics (means, standard deviations) for expert scores and the computation of Intraclass Correlation Coefficients (ICC) using the two-way mixed effects model ICC(3,1) to assess absolute agreement between the two raters. ICC interpretation followed established thresholds to flag competencies or items with acceptable versus inadequate inter-rater reliability (Koo & Li, 2016). Instead of formal qualitative coding, the quantitative findings were juxtaposed with the general student observations. This informal triangulation helped to contextualize convergent and divergent patterns in the expert ratings, ultimately guiding recommendations for future rubric refinement and rater calibration.

3. DISCUSSION

Across the two expert raters, the combined mean scores (scale 1–5) indicate that the task set was judged strongest for STEM Knowledge & Application (mean = 4.75), followed by Reasoning & Decision Making (mean = 4.17) and Information & Data Processing (mean = 4.17). Critical Thinking (mean = 3.83) and Problem & Conflict Solving (mean = 3.67) received lower average ratings. These combined means suggest that experts consistently perceived the items as particularly effective for eliciting applied STEM and data-processing skills (see Table 1).

Table 1. Rater perspective by competency

Competency (domain)	**Expert 1 mean**	**Expert 2 mean**	**Combined mean**	**ICC(3,1) (point estimate)**
Reasoning & Decision Making	4.00	4.33	4.17	0.00006
Problem & Conflict Solving	4.00	3.33	3.67	0.750
Critical Thinking	3.83	3.83	3.83	0.000

Information & Data Processing	4.17	4.17	4.17	0.778
STEM Knowledge & Application	4.83	4.67	4.75	0.615

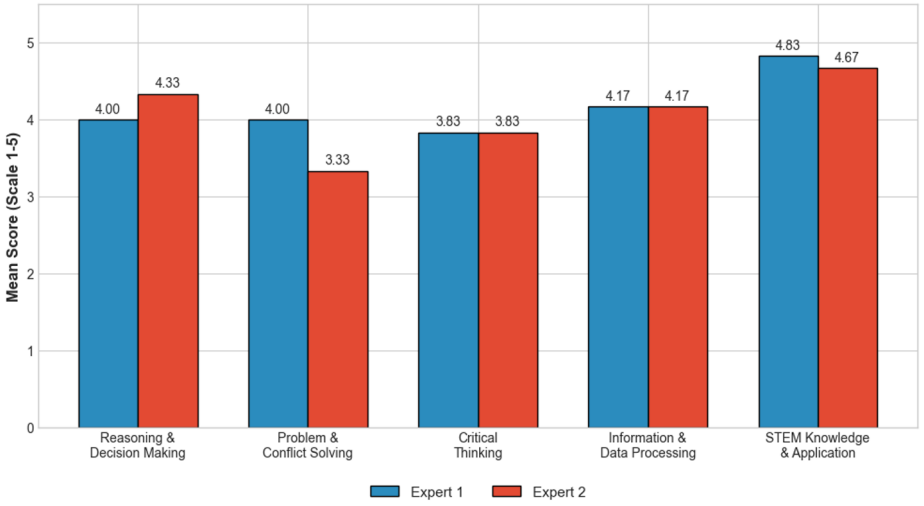

Figure 1. Comparison of Expert Rating by Competency

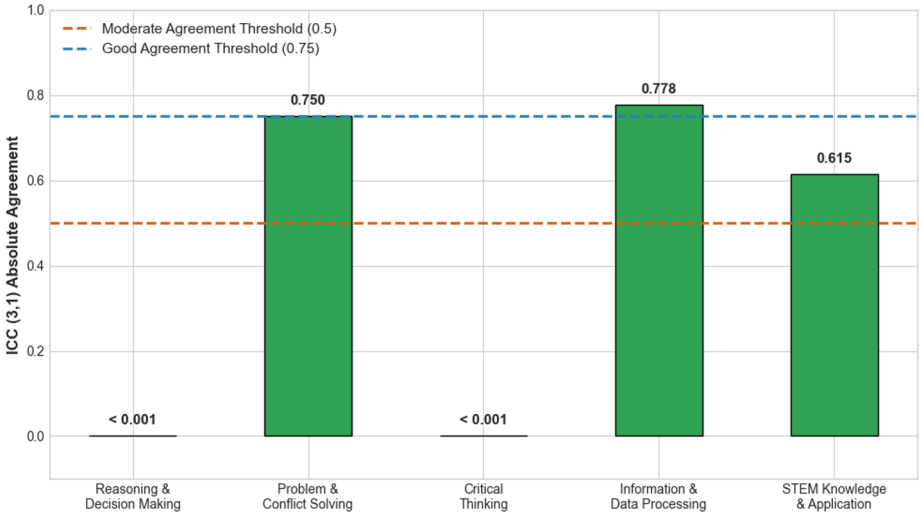

Figure 2. Inter-rater Reliability (ICC) by Competency Domain

A closer look at rater behavior reveals a systematic tendency: Expert 2 generally awarded higher scores for Information & Data Processing (≈ 4.83) and STEM (≈ 4.67) than Expert 1, who was more conservative on those domains (Information ≈ 3.33;

STEM ≈ 3.83). This difference accounts for some of the variability in competency means and highlights the need for rater calibration when assessing applied competencies.

The item-level competency profiles and the experts' chosen dominant competencies are visually summarized in Figure 3.

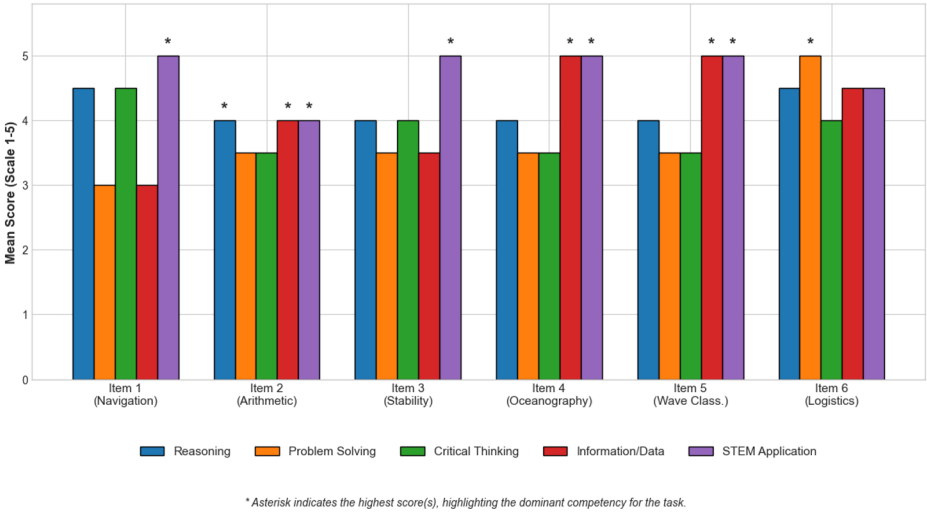

Figure 3 the mean scores for each of the five competencies across the six maritime tasks

As depicted in Figure 3, data-rich calculation tasks (Items 4 and 5) robustly elicit Information & Data Processing and STEM application. Conversely, operational decision tasks (such as the logistics scenario in Item 6) primarily elicit problem-solving competencies. Meanwhile, tasks centered on navigation and stability (Items 1 and 3) effectively bridge STEM knowledge with reasoning skills.

Inter-rater reliability was evaluated with ICC(3,1) (two-way mixed, absolute agreement). By competency, the ICC point estimates were: Information & Data Processing = 0.778 (good agreement), Problem & Conflict Solving = 0.750 (borderline good/moderate), STEM Knowledge & Application = 0.615 (moderate), Reasoning & Decision Making ≈ 0.00006 (negligible), and Critical Thinking = 0.000 (negligible). In short, experts achieved reliable agreement on measurable data-processing behaviors and reasonably on problem-solving/STEM, but they did not reach consensus on reasoning and critical thinking as operationalized in the rubric.

ICC by item also varied: Item 1 ICC = 0.795 (good); Item 2 ICC ≈ 0.00006 (negligible); Item 3 ICC = 0.667 (moderate); Item 4 ICC = 0.704 (moderate/good); Item 5 ICC = 0.704 (moderate/good); Item 6 ICC = 0.000 (negligible). These item-level ICCs support the experts' confidence ratings: Items 4–5 and Item 1 produced the strongest agreement, while Items 2 and 6 require rubric clarification or rater calibration before their ratings can be treated as highly reliable.

Student Evidence General student observations provided practical triangulation. Students' general approaches clearly demonstrated instances of STEM application (use

of formulas and domain reasoning) and Information & Data Processing (data extraction and unit conversion) on oceanography and engineering items. Conversely, evidence of Critical Thinking, explicit evaluation of assumptions, discussion of precision, and reflective justification, was weak in this small student sample. Overall, general observations indicate that maritime numeracy questions effectively brought out STEM Knowledge & Application and Information & Data Processing, while the Critical Thinking aspect remains an area needing stronger prompt design to encourage the evaluation of assumptions and implications. Thus, these practical insights corroborate the strongest expert claims and point to a clear area for improvement in item design.

General student observations provided practical triangulation. Students' general approaches clearly demonstrated instances of STEM application (use of formulas and domain reasoning) and Information & Data Processing (data extraction and unit conversion) on oceanography and engineering items. Conversely, evidence of Critical Thinking, explicit evaluation of assumptions, discussion of precision, and reflective justification, was weak in this small pilot sample. Overall, general observations indicate that maritime numeracy questions effectively brought out STEM Knowledge & Application and Information & Data Processing, while the Critical Thinking aspect remains an area needing stronger prompt design to encourage the evaluation of assumptions, implications, and answer precision. Thus, these practical insights corroborate the strongest expert claims and point to a clear area for improvement in item design.

These findings align closely with prior research. The prominence of STEM and Reasoning in expert judgments resonates with rankings of critical competencies for future seafarers (Ceylani et al., 2022) and with literature positioning mathematical literacy as central to maritime operations (Novitasari et al., 2024; Stanivuk et al., 2017). Moreover, the pattern of higher agreement on concrete, observable behaviors (data manipulation, formula use, stepwise decomposition) and lower agreement on abstract constructs (reasoning, critical thinking) mirrors established assessment challenges (Koo & Li, 2016; Niss & Hojgaard, 2011). Consequently, improving the operational definitions, providing exemplar answers, and conducting rater calibration are necessary steps to strengthen reliability for assessing reasoning and critical thinking.

Finally, the finding that workplace-relevant logistics problems strongly elicit problem-solving (Item 6) supports claims that authentic, domain-specific tasks are effective for fostering practical mathematical competencies in maritime education (Stanivuk et al., 2017). In light of recent calls to update MET/STCW frameworks to include data literacy and higher-order cognition (Ghosh et al., 2024; Belabyad et al., 2025), this study's descriptive analytical approach, combining task design, expert competency mapping, ICC diagnostics, and practical student observations, offers a viable model for aligning assessment design with evolving industry needs.

4. CONCLUSION

Overall, the developed maritime-scientific numeracy tasks function effectively as a dual-purpose tool: they serve both as assessment items and as learning activities that connect mathematics to maritime practice. However, this study demonstrates that maritime-scientific numeracy tasks are most effective at eliciting STEM Knowledge & Application and Information & Data Processing, as confirmed by both expert ratings

and general student observations. Evidence for reliably capturing Reasoning and Critical Thinking remains weak due to low inter-rater agreement. The strongest agreement was observed for data-oriented items (ICC ≈ 0.78), whereas abstract reasoning and critical-thinking indicators require clearer operational definitions and rater calibration.

Practically, the current item set can be readily utilized by Maritime Education and Training (MET) instructors to teach and assess applied STEM and data skills. However, before using these items to make high-stakes judgments about reasoning or critical thinking, the rubric should be refined, exemplar answers provided, and raters trained. A limitation of this study is the reliance on informal student observations; future research should involve formal quantitative and qualitative assessment of a larger student sample, alongside additional raters, to further validate these findings.

References

Belabyad, M., Kontovas, C., Pyne, R., Shi, W., Li, N., Szwed, P., & Chang, C. H. (2025). The human element in autonomous shipping: a study on skills and competency requirements. *WMU Journal of Maritime Affairs*. https://doi.org/10.1007/s13437-025-00366-9

Ceylani, E., Kolçak, Ç., & Solmaz, M. S. (2022). A Ranking of Critical Competencies for Future Seafarers in the Scope of Digital Transformation. *Proceedings of the International Association of Maritime Universities Conference*, *2022-October*.

Fachrudin, A. D., Rahmawati, T. D., & Novitasari. (2025). *Desain Soal Numerasi Konteks Saintifik dengan Integrasi Konsep Maritim*. *14*(2), 565–579. https://doi.org/10.26740/mathedunesa.v14n2.p565-579

Fachrudin, A. D., Widadah, S., & Kusumawati, I. B. (2019). Pre-service mathematics teachers' knowledge, beliefs, and attitude toward using PISA-based problem in mathematics education. *Journal of Physics: Conference Series*, *1200*(1). https://doi.org/10.1088/1742-6596/1200/1/012013

Ghosh, S., Emad, G. R., & Ravi, A. (2024). Investigating the characteristics of skills and competency frameworks through a systematic literature review: a feasibility study to revise the STCW Code for seafarer training. *Australian Journal of Maritime and Ocean Affairs*. https://doi.org/10.1080/18366503.2024.2374606

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2). https://doi.org/10.1016/j.jcm.2016.02.012

Mycroft, K., & Sriraman, B. (2020). Mathematics in the Maritime. In *Handbook of the Mathematics of the Arts and Sciences*. https://doi.org/10.1007/978-3-319-70658-0_137-1

Niss, M., & Højgaard, T. (2011). Competencies and Mathematical Learning Ideas and inspiration for the development of mathematics teaching and learning in Denmark. In *IMFUFA tekst, Roskilde University, Denmark* (Number 485).

Novitasari, N., Agustina, E. N. S., Fachrudin, A. D., & Irfan, S. (2024). Mathematical Literacy as a Key Competency for Maritime Education: Systematic Review of

Cadet Competencies. *VYGOTSKY*, *6*(2), 131–140. https://doi.org/10.30736/voj.v6i2.1112

Stanivuk, T., Galić, S., & Bojanić, M. (2017). Mathematics as a science and marine activity follow each other throughout history. *Transactions on Maritime Science*, *6*(1). https://doi.org/10.7225/toms.v06.n01.006

Zaura, B., Fatimah, S., Hidayat, M., & Harnita, F. (2025). Development of Realistic Mathematics Education-based Numeracy Test Instrument with Maritime Context for Junior High School Students in Banda Aceh, Indonesia. *Jurnal Matematika Kreatif-Inovatif*, *16*(1), 44–61.