Military Staff Study as a Tool for Language Policy Design: A Case of Establishing the RTN's English Tests

Jitpicha Jarayapun
Royal Thai Navy, Thailand

ABSTRACT

A staff study (SS) is an administrative tool used for decision-making, involving comprehensive data gathering and analysis to resolve issues. Particularly crucial in military settings where classified information limits resources, staff studies also demonstrate their value in interdisciplinary academic fields. This article emphasizes the significance of SS in education policy by exploring their application in developing English testing within the Royal Thai Navy (RTN). While staff studies share similarities with applied research in language education, their primary emphasis is on applying knowledge to solve problems in specific contexts. Staff study principles provide a logical framework for navigating contextual variables in problem-solving, rather than expanding knowledge. Overall, SSs offer innovative approaches to addressing unique challenges across various research contexts.

Keywords: English language policy, language testing development, military staff study

INTRODUCTION

The overall English proficiency of the Thai population was ranked 101 out of 113 nations in 2023; this places Thailand in the very low category (EF, 2024). Many analyses indicate that Thai people's attitude toward their own language reflects Thai as the language of national identity. The Thai language is regarded as a matter of pride for the only country in Southeast Asia which was never colonized by any European power. Unsurprisingly, most Thai people rarely use English in their daily

life, not even in their work-related communication. However, a new trend of English requirements for work related tasks has greatly increased their awareness of English in a globalized economy. Hayes (2014) stated that English proficiency is perceived as "a tool for personal economic advancement"; nonetheless, the Thai language is still unchallenged and unthreatened by English. The sense of preserving a long-treasured language does not assist with attempts to provide English learning environments. Thus, it is a significant challenge for the Thai government to support their people to develop English proficiency when most Thais prefer using Thai as the language of their national pride.

The Royal Thai Navy (RTN), a governmental organization of approximately 40,000 active personnel categorized into 23 professional fields, has realized the critical needs of improving their personnel's English proficiency to effectively collaborate with military and civilian staffs from other countries. International cooperation is vital for achieving mutual goals such as security protection, humanitarian operations, and disaster relief. To fulfill the RTN's English demand, various policies, English programs, and learning facilities have been implemented to improve the English skills of naval personnel. Recently, the RTN released the Master Plan for the Development of Naval Personnel Management System for 20-Year Strategic Plan (B.E. 2560–2579) (Royal Thai Navy, 2017), which emphasizes the importance of improving the English proficiency of RTN personnel to enable effective collaboration with staff from other nations, as well as, strengthening international partnerships and achieving mutual goals. The objectives of this plan underscores the Thai Navy's goals for enhancing personnel's English skills of its personnel. These objectives include supporting the ability to participate effectively in joint and combined operations and training, international seminars, discussions, and meetings.

The use of language tests as "instruments of policy" has been widely acknowledged, providing a practical mechanism for policy implementation (Shohamy, 2007). Similarly, within the RTN, English language testing holds strategic significance in the pursuit of improving officers' English proficiency. Testing plays a role in identifying, classifying, and the assignment of officers into positions involving international functions. The evaluation has been integrated as a crucial aspect of Human Resource Management. One prominent language test used by the Royal Thai Navy is the American Language Course Placement Test (ALCPT), developed by the Defense Language Institute English Language Center (DLIELC). This test serves as an evaluation tool for assessing officers' English skills and determining their qualifications. Widely recognized as a certificate placement test, it was initially designed to recruit candidates for foreign military and US government-sponsored English Language Training Programs (ELTP) worldwide (DLIELC, 2024).

Language policies within the RTN have included the implementation of the ALCPT. These policies, including the RTN strategic plan FY 2006-2015, were

designed to position the RTN as a leading force in Southeast Asia. To align with the RTN's strategic vision for 2025, Human Resource Management made it mandatory for officers in higher ranks of Captain or full Captain to achieve specific ALCPT scores. Officers ranked from CDR to CAPT were required to obtain scores between 74 and 79, while full captains needed scores between 80 and 84. These score requirements were put in place to ensure the readiness of RTN personnel's English proficiency in line with the RTN's long-term goals (Royal Thai Navy, 2019).

Despite the use of ALCPT scores as a measure of English proficiency, there have been limited positive outcomes in improving the RTN's English language skills. Concerns have been raised regarding officers who obtained high scores on the test but displayed incompetence in effectively using English during joint military trainings. This discrepancy between test scores and practical English language proficiency among RTN personnel has led to dissatisfaction with the outcomes of using ALCPT as a policy tool. In an effort to address these challenges and enhance English proficiency among RTN personnel, the RTN implemented the English language policy based on the approach of the Common European Framework of Reference for Languages (CEFR) in FY 2021. This led to the development of the "Royal Thai Navy English Proficiency Test (NEPT) Prototyping" project by the Naval Education Department (NED) (Jarayapun, 2021). The NEPT aims to establish a standard-based assessment that extends officers' English proficiency, provides a classification of language proficiency aligned with the CEFR, and ensures a fair and objective accountability system within the RTN. However, there is a need for significant development in the NEPT, particularly in the implementation of a formal standard setting procedure. This procedure should involve clearly defining and marking the cut-off scores as boundaries between six benchmarks, as outlined by the Common European Framework of Reference for Languages (CEFR).

Language testing has been widely used to support English language development in policy implementation (Shohamy, 2007; Shohamy, 2014; Menken, 2008). However, it can either promote or restrict language teaching and learning (Wiley & Garcia, 2016). Therefore, understanding the contexts and circumstances is crucial because testing is a critical component for language development.

To address the RTN's need for a strategic English learning plan, a military staff study titled "Developing the Royal Thai Navy's English Test as a Strategic Tool to Improve English Skills for International Collaboration" was conducted. This study applied the staff study principle to solve the research problem, providing insights into the RTN's organizational circumstances through language education research. Using a descriptive research design and document analysis, the study explored the staff study principle in developing language tests.

The author explores two key questions:

1) What are the characteristics of superior English tests that can enhance RTN personnel's language proficiency and achieve the navy's aspirations?

2) How does a staff study determine its effectiveness in solving practical problems, and how does it differ from methods used in applied research?

To address these questions comprehensively, the study examines the concept of military staff study and its specific application in developing English tests within the RTN.

LITERATURE REVIEW: STAFF STUDY PRINCIPLE

The "Staff Study" (SS) is a methodological process conducted by a staff officer to examine an issue in depth. It involves comprehensive data collection and analysis to identify alternative courses of action (COAs), which are then refined into an optimal solution and presented for approval. In the Royal Thai Navy (RTN), staff work follows seven sequential steps, classified into three primary stages: (1) problem identifying, (2) data gathering, and (3) solution formulation (Institute of Advanced Naval Studies, 2016).

Figure 1 Staff Study Principle
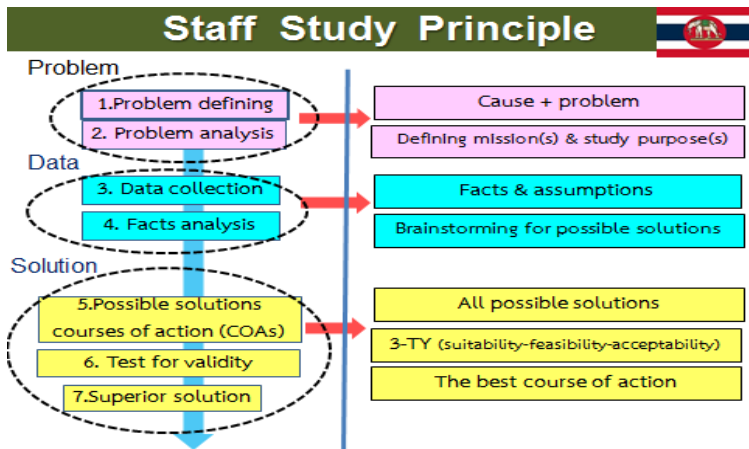
(Adapted from Institute of Advanced Naval Studies [PowerPoint slides], 2016)

1. Problems Identifying: Defining and Analyzing Problems

The initial stage of problem identification in SS involves understanding issues to gather relevant data to identify optimal solutions. Problems are categorized into three types: a) urgent issues requiring immediate attention, b) preventive problems, and c) developmental problems. In this study, the SS pertains to developmental problems as it aims to improve the quality of RTN's EFL test,

enhancing work performance from "acceptable" to "superior". Relevant information is thoroughly gathered from policies, measurements, orders, and annual plans to systemically examine underlying causes and effects.

Defining the problem in a staff study is similar to a "problem statement" within a given topic in research (Ary et al., 2010; Mills & Gay, 2016; Creswell & Creswell, 2018). While research problem guides research plan and serves as a justification for the educational process, problem defining in SS is merely facilitating rational decision-making for a specific issue. Unlike a research hypothesis that predicts outcomes, the SS hypothesis serves as an assumption regarding factual relationships. It is optionally applied when information is incomplete or lacks clear evidence before conducting the study analysis; thereby, assisting rational decision-making for specific issues.

2. Data Collection and Facts Analysis

Data collection and facts analysis in the staff study (SS) resembles "literature reviews" in traditional research design frameworks (Sahni & Sinha, 2016; Snilstveit et al., 2012; Machi & McEvoy, 2016). The primary objective is to systematically gather and organize all relevant facts related to the identified problem. These facts are analyzed from various perspectives, considering all sides of an argument, until logical (inductive) conclusions can be derived. The SS supports informed decision-making rather than highlighting research gaps like traditional research.

3. Decision Process for Solution

This phase bears similarities to a "research plan," systematically evaluates all study dimensions, providing guidance for its implementation (Mills & Gay, 2016). Decision-making tools used in RTN staff study include Validity Analysis (VA) and the decision matrix.

3.1 Validity Analysis: Exploring Possible Courses of Action (COAs)

Validity Analysis (VA) identifies diverse Courses of Action (COAs) through brainstorming and comprehensive fact examination. Each COA must be distinct, offering benefits aligned with study objectives. It must be feasible within available resources within a reasonable timeframe, and manageable regarding associated risks.

To determine the potential COAs in Validity Analysis, staff officers use "3TY filtering" to evaluate COAs based on:

a. Suitability: alignment with SS objectives to fulfill the intended goals.

b. Feasibility: implementation with available resources, commonly referred as 8M+T aspects (man, money, material, management, methods, machines, market, messages or information, and time).

c. Acceptability: valuable outcomes despite potential obstacles.

COAs are refined until they meet these criteria, ensuring they are suitable, feasible, and acceptable.

Figure 2 3-TY Filtering in Staff Study
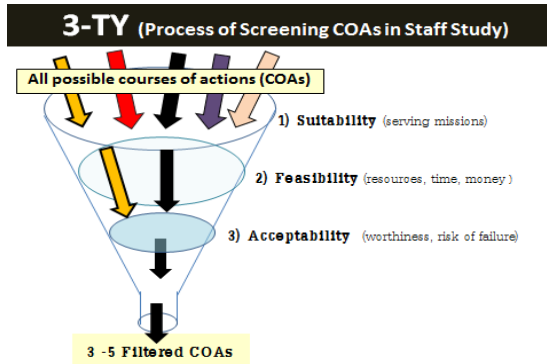
(Adapted from Institute of Advanced Naval Studies [PowerPoint slides], 2016)

3.2 Superior Solution

The "Superior Solution" stage uses quantitative methods such as Analytical Hierarchy Process (AHP) and Maximin-Minimax. Within the RTN, the decision matrix or selection matrix is commonly employed to compare problem-solving potential; thus, identifying the most optimal choice among selected COAs.

The SS principle shares similarities with applied research, aiming to identify specific and promising practices for solving educational problems (Mills & Gay, 2016). The decision-making process in SS serves as a well-structured "research plan" and methodology for investigating optimal solutions to problems. However, while critical to the research process, the SS is not research itself but a methodically devised research design that has yet to be implemented.

STAFF STUDY EXPLANATION

This section demonstrates the practical implementation of validity analysis and the decision matrix to identify the optimal English test for the RTN. The primary goal is to design an English test that will yield the greatest positive outcomes in terms of language acquisition for RTN personnel. The intention is to equip officers with the ability to effectively communicate in English with naval staff from partner

nations within regional and global networks, thereby enhancing the overall professionalism of the RTN. The following analytical description outlines the evaluation of each Course of Action (COA) using the 3-TY filtering approach within the SS framework.

1. Possible Courses of Action (COAs)

Understanding the efficacy of the staff study principle involves thorough investigation of all relevant facts to identify viable solutions. The following analysis outlines the application of the 3-TY filtering approach.

1.1 Royal Thai Navy English's English Language Test (RTN-ELT)

The Royal Thai Navy English's English Language Test (RTN-ELT) resembles the American Language Course Placement Test (ALCPT), a standard test for evaluating English language proficiency of local personnel seeking positions on overseas US military installations (DLIELC, 2024). It features 100 multiple-choice questions with four answer options covering 66 listening and 34 reading items. The RTN-ELT focuses on current global topics and local-international news stories, distinguishing it from the ALCPT.

Designed to evaluate RTN personnel's English proficiency for international work, the RTN-ELT meets criteria of suitability, feasibility, and acceptability. The test encourages awareness of global affairs among RTN personnel, emphasizing relevant information in global contexts. The Royal Thai Navy Language Center (RTNLC) has well-trained and experienced instructors with expertise in designing English tests to ensure the test's effectiveness and validity. Moreover, a team of RTN's English teachers from diverse backgrounds are also involved to support the inclusion of international English accents.

The estimated cost of establishing and maintaining RTN-ELT is manageable within the Thai Naval Educational Department's budget, totaling approximately 92,000 THB (2,300 USD) annually and 460,000 THB (11,500 USD) over five years. Enhancing test effectiveness involves rigorous analysis by experienced English teachers at RTNLC focusing on factors like reliability, difficulty, and item discrimination to ensure test quality and validity.

1.2 Royal Thai Navy's English Proficiency Test (NEPT) in Revision

The Royal Thai Navy's English Proficiency Test (NEPT) underwent significant development through NEPT Prototyping in 2021 (Jarayapun, 2021), adopting question types from the globally recognized TOEIC test, known for its accurate assessment of English-language listening and reading skills (ETS, 2024). The revised NEPT will maintain 30 questions and a 35-minute time allocation,

continuing to be administered online to RTN officers with the cooperation of the Naval Educational Department and the Naval Communications Department.

To enhance NEPT quality and align it with the Common European Framework of Reference for Languages (CEFR), a standard setting procedure (Council of Europe, 2009; Livingston & Zieky, 1982) will validate its scores and equate them with the CEFR standard scales. Statisticians within RTN will ensure the test's validity and alignment with CEFR standards.

The revised NEPT fulfills the criteria of the 3-TY (suitability, acceptability, and feasibility), expected to provide comprehensive and comparable data aligned with the CEFR standards. The test content is tailored to assess language proficiency essential for international work. The estimated cost of revising or creating a new NEPT is 963,000 THB (24,075 USD). This includes the TOEIC test fee for 396 participants in the NEPT-TOEIC alignment study and the exam administration over five years. These funds can be managed within the budget allocated to the RTN Test Development Project, pending RTN approval.

1.3 English for Specific Purposes Test (ESP Tests)

The design of ESP tests for the RTN will be based on the principles of English for Specific Purposes (ESP) assessment (Douglas, 2013; Hutchinson & Waters, 1987; Sabieh, 2018), tailored to the 23 job families within the organization. According to Dudley-Evans and St John (1998), ESP tests should involve specialists who understand the language requirements of their fields. Therefore, the development of ESP tests will include collaboration between language teachers and stakeholders from the 23 job families within the RTN who possess extensive competence and experience in their respective areas. Utilizing internal RTN resources, the process will begin with a needs analysis while considering the assessment purposes and the context of language use within specific job roles (Paltridge & Starfield, 2013). The Naval Educational Department (NED) can host meetings and conduct the necessary needs survey to facilitate this process. The proposed budget for test development, estimated at approximately 1,268,000 THB (31,700 USD), is reasonable and subject to approval by the Director General of the NED.

ESP tests offer a high return on investment for the RTN, being reusable as work-related tasks remain consistent and applicable to various grammatical contexts over time. By distributing topics and details in advance for test preparation, concerns about cheating within the RTN can be minimized. Additionally, the needs analysis conducted for ESP test development can inform curriculum design, ensuring that educational programs cater to the specific needs of RTN officers. ESP tests serve as a problem-solving tool rather than solely a measurement tool for human resource management. This strategic approach supports English learners in all levels within the RTN. Notably, the constructive

impact of ESP tests on learners contributes to the concept of "beneficial backwash," where testing positively influences the learning and teaching process (Bailey, 1996; Messick, 1996; Hughes, 2003; Green, 2013). Although the development of ESP tests requires significant time and management, the long-lasting benefits are advantageous.

2. Superior Tests: Utilizing the Decision Matrix Approach

The Decision Matrix determines the most suitable test among various courses of action (COAs) through two steps: (a) describing criteria and (b) assessing COAs using the decision matrix. Each COA is scored from 1 (lowest) to 5 (highest) based on its anticipated impact. Clear descriptions ensure impartiality in scoring. Relevant criteria are established by considering "effectiveness," "efficiency," and "economy" (the "3E" principle). The table below presents the descriptions for six criteria used in the decision matrix:

Table 1 Criterion Description

Effectiveness		
Criteria 1	Ability to Serve RTN's Purposes (Needs for Work)	
Scales		Quality Description
5	Highest	Tests fully ensure that learners can effectively communicate in working contexts aligned with the RTN's needs
4	High	Tests can somewhat encourage learners to communicate in working contexts aligned with the RTN's needs
3	Moderate	Tests cover general topics that are likely useful for work aligned with the RTN's needs.
2	Low	Tests may or may not support the RTN's needs for work.
1	Lowest	Tests are clearly irrelevant to the RTN's needs for work.
Criteria 2	Test Credibility	
5	Highest	Tests demonstrate evidence of a high-quality language testing system that accurately interprets individuals' language abilities. They are approved for their reliability, validity, and alignment with recognized fairness and validity criteria or accredited tests.
4	High	Tests somewhat demonstrate evidence of a high-quality language testing system with the potential for accurate

interpretation of individuals' language abilities based on global standards.

3	Moderate	Tests can be used to assess individuals' language abilities but raise doubts about their accuracy.
2	Low	Test results are unlikely to be credible in measuring individuals' language abilities accurately. Limited validity processes are conducted to verify test quality.
1	Lowest	Test results cannot be trusted as a valid assessment of individuals' language abilities due to a lack of quality verification.

Criteria 3 Reusability

5	Highest	Test questions are fully reusable for assessing the language abilities of individual RTN personnel with fairness and validity.
4	High	Test questions can be reused after many years of retention to assess the language abilities of individual RTN personnel, but it may decrease test fairness.
3	Moderate	Test questions can somewhat be reused after many years of retention to assess the language abilities of individual RTN personnel, but it may decrease test fairness.
2	Low	Test questions should not be reused to assess the language abilities of individual RTN personnel even after many years of retention, as they may become outdated, useless, or unfair.
1	Lowest	Test questions are not recommended for reuse. They lack validity in supporting learning or communication skills for work.

Efficiency

Criteria 4 Manageability

5	Highest	Establishing the test can be smoothly controlled and supervised within a short operating period (within 3 months) with limited cooperation from specialized expertise.
4	High	Establishing the test can be controlled and supervised with relative ease, requiring a moderate operating period (4-6 months) and some cooperation from specialized expertise.
3	Moderate	Establishing the test can be done with careful planning, control, and supervision, requiring a sufficient operating

		period (6-12 months) involving one or more groups of specialized expertise.
2	Low	Establishing the test is quite challenging, requiring meticulous planning, complete control, and supervision from all relevant sectors involving specialized expertise over a period of 1-2 years.
1	Lowest	Establishing the test is extremely challenging, necessitating exceptional planning, complete control, and supervision from all relevant sectors, both internally and externally, involving specialized expertise for more than 2 years or indefinitely.

Criteria 5	Positive impacts on self-learning and language development

5	Highest	Tests fully motivate learners to actively develop their English skills for work. The test content and instructions effectively guide test takers on how to improve their English proficiency. Test takers gain confidence in their English abilities upon achieving desired test scores, indicating desirable competency in English proficiency.
4	High	Tests somewhat motivate test takers to develop their English skills for work, as they address relevant topics that imply ways to improve English proficiency.
3	Moderate	Tests may create awareness among takers regarding the need for English improvement, but they lack explicit guidance on how to develop better English skills without additional suggestions.
2	Low	Tests may increase test takers' motivation to pass the test, but they do not provide guidance on how to improve English for work effectively. It is challenging to determine how to enhance English skills solely based on the test.
1	Lowest	Test takers feel uncomfortable taking the tests, as they fail to recognize the relevance of the tests to their ability to communicate in English for work.

Economy

Criteria 6	Budget (calculated for five years; 40 THB = 1 USD)

5	Highest	0 – 100,000 THB, or approximately 0 – 2,500 USD
4	High	100,001 – 600,000 THB, or approximately 2,501 – 15,000 USD

3	Moderate	600,001 – 1,100,000, or approximately 15,001 – 27,500 USD
2	Low	1,100,001 – 1,600,000 THB, or approximately 27,501 – 40,000 USD
1	Lowest	Over 1,600,000 THB, or over 40,000 USD

The Decision Matrix compares COAs based on six criteria, shown in the Decision-Making Spider Web Chart (See Figure 3).

Each criterion in the Decision Matrix is weighted for importance, considering both qualitative and quantitative aspects. The ESP tests (COA3) emerge as the superior solution, scoring highest in three key areas: serving RTN's purposes, reusability, and positive impacts on self-learning and language development (see Table 2).

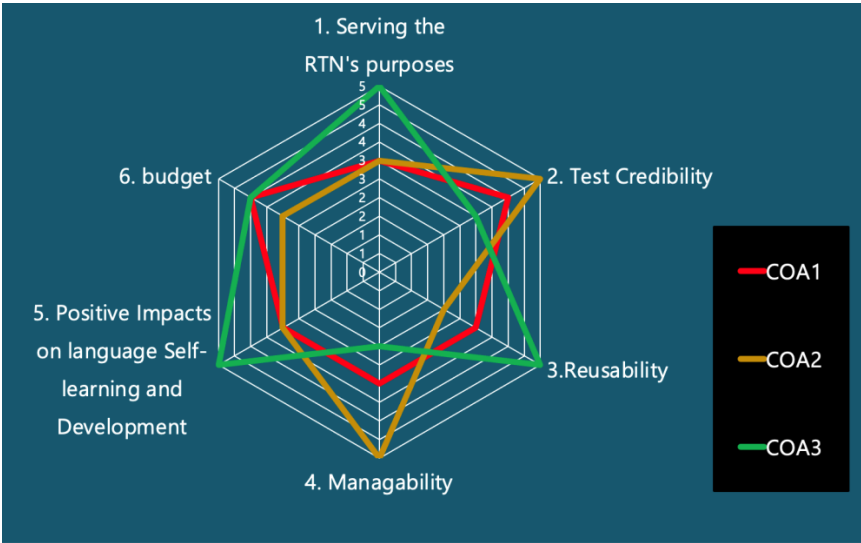**Figure 3 Spider Web Chart for Selecting the Superior Test**

Table 2 Decision matrix to select the superior solution

Criteria	Weights (W)	COA 1		COA 2		COA 3	
		Scores	S x W	Scores	S x W	Scores	S x W
1. Serving the RTN's purposes	30	3	90	3	90	5	150
2. Test Credibility	20	4	80	5	100	3	60
3. Reusability	10	3	30	2	20	5	50

4. Manageability	10	3	30	5	50	3	30
5. Positive impacts on Language development	20	3	60	3	60	5	100
6. Budget	10	4	40	3	30	2	20
Total	100		330		350		410

DISCUSSION: THE SUPERIOR ENGLISH TEST

The analysis has determined that the use of ESP tests is the best solution to align with the RTN's strategic vision and fulfill their objectives. Exploring ESP concepts and researching ESP learning and teaching revealed several additional benefits for the RTN as follows:

1. ESP testing serves as a powerful tool for educational administration. The development of ESP competency and professional expertise can effectively support lifelong English usage. Tailoring English for work to the specific needs of each of the RTN's 23 job families fosters a sense of professionalism within the organization.

2. Contextualizing language and learning is crucial to support the development of ESP competency and prepare personnel for ESP tests (Gollin-Kies et al., 2015). This approach encourages the practical use of English in various workplace settings. By increasing awareness and relevance of English in their work environment, personnel will embrace English as a language integral to their professional responsibilities.

3. Achieving success in ESP tests ensures test-takers' ability to effectively use English in their work environment. This accomplishment is facilitated by effective management of social factors, trust levels, support, and coordination. Consequently, increased interaction between organizational leaders and subordinates can be expected.

4. The development of ESP tests requires careful consideration of knowledge, work achievements, and socio-cultural aspects. Although it may be challenging to design efficient ESP tests, the effort invested in their development is worthwhile not only for managerial support but also for the overall language curriculum design within the RTN.

By implementing ESP tests, the RTN can capitalize on the multitude of benefits they offer, ranging from supporting lifelong English usage to fostering professionalism and improving communication within the organization. Furthermore, the careful design and development of ESP tests will have a positive impact on the RTN's language curriculum and overall language learning environment.

CONCLUSIONS AND IMPLICATIONS

This study exemplified the implementation of staff study (SS) in the development of English tests for the RTN. The relationship between SS and research was explored within the context of language testing and education.

1. SS relies on the collection and analysis of scientifically obtained data. While SS is not primarily oriented towards research outcomes, it can serve as an initial step, particularly in the realm of "action research" (Nicodemus & Swabey, 2015). Additionally, SS can provide a foundation for the development of method-specific implementations aimed at enhancing the quality of large organizations.

2. The sensible decision-making approach of 3-TY filtering can be applied to various decision-making processes. The Decision Matrix offers a logical framework for efficiently comparing potential solutions. Describing decision scores and weighting various criteria can shed light on the advantages and limitations of each course of action. By carefully considering SS, risks, errors, and uncertainties associated with decision-making can be managed.

3. The results of SS are adaptable to specific focuses and study purposes. For instance, when evaluating intercultural competency in global networks, different COAs can be formulated. The outcomes of SS bear resemblance to modeling key research variables within socially diverse contexts.

4. Like other types of research, SS relies on the expertise of the investigators. Although the SS principle aims to minimize staff preferences, the selection of facts and analysis is carried out by staff officers who may be influenced by their own inclinations. To mitigate risks, biases, and errors in decision-making, staff officers need to be mindful of their biases and recognize how such biases can potentially limit their decisions. Sharing facts and ideas with a group of experienced staff officers can help mitigate biases. Furthermore, involvement of scholars can enhance the reliability of the study.

5. Lastly, effective decision-making requires careful deliberation and sufficient time to explore the facts. However, decision-making skills can be honed through training, which can ultimately reduce the time required for decision-making processes.

In conclusion, while research is commonly considered the standard approach to inquiry, it is important to recognize that there are multiple perspectives and methodologies available for harnessing existing knowledge and driving advancements in language education. One such approach is Staff Study (SS), which provides an alternative means of understanding the world and addressing a diverse range of challenges in EFL testing. By employing the SS procedure, it is possible to engage in empirical sense-making and establish rigorous standards that are vital for effective language policy design and informed decision-making in practical contexts.

By embracing SS as a complementary option alongside traditional research methodologies, the toolkit is expanded, and perspectives are broadened, leading to a more comprehensive and holistic understanding of language education. SS offers a logical framework that enables us to make sense of complex issues and develop practical solutions. The establishment of rigorous standards through the SS procedure ensures a solid foundation for language policy design and facilitates informed decision-making, empowering stakeholders in the field of language education to enact positive changes and achieve desired outcomes. Therefore, considering the various ways of looking at the same subject matter, incorporating SS as an additional approach can significantly contribute to the ongoing development and improvement of language education.

REFERENCES

Ary, D., Jacobs, L. C., Sorensen, C., & Asghar, R. (2010). *Introduction to research in education*. Wadsworth Cengage Learning.

Bailey, K. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing, 13*(3), 257-279.https://doi.org/10/1177/026553229601300303

Council of Europe. (2009). *Relating language examinations to the 'Common European Framework of Reference for Languages: Learning, teaching assessment' (CEFR), a manual.* Council of Europe. http://www.coe.int/en/web/common-european-framework-languages/relating-examinations-to-the-cefr

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE.

DLIELC. (2024, June 13). *2024-2025 Defense English Language Program: Course catalog and security cooperation handbook*. Defense Language Institute English Language Center. https://www.dlielc.edu/prod/Catalog. pdf

Douglas, D. (2013). ESP and assessment. In B. Paltridge, & S. Starfield (Eds.), *The handbook of English for Specific Purposes* (pp. 367-383). Wiley-Blackwell.

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.

EF. (2024). *EF English Proficiency Index*. EF. https://www.ef.co.th/epi/about-epi/#test-takers

ETS. (2024). *Educational Testing Service*. ETS. https://www.ets.org/toeic

Gollin-Kies, S., Hall, D. R., & Moore, S. (2015). *Language for specific purposes*. Palgrave Macmillan.

Green, A. (2013). Washback in language assessment. *International Journal of English Studies, 13*(2), 39-51. https://doi.org/10.6018/ijes.13.2.185891

Hayes, D. (2014). The value of learning English in Thailand and its impact on Thai: Perspectives from university students. *Asia Pacific Journal of Education, 36*(1), 71-91. https://doi.org/10.1080/02188791.2014.924390

Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learner-centered approach*. Cambridge University Press.

Institute of Advanced Naval Studies (2016). เอกสารอ้างอิงของกองทัพเรือหมายเลข 8302 [Royal Thai Navy Reference No.8302]. Nakorn Pathom, Thailand: Royal Thai Naval Education Department.

Jarayapun, J. (2021, September). CEFR กับการสร้างข้อสอบต้นแบบ NEPT [CEFR and the Royal Thai Navy's English Proficiency Test Prototype]. *นาวิกศาสตร์, 104*(9), 22-38. https://www.rtni.navy.mi.th/upload/pdf/Article2564/055.pdf

Livingston, S.A., & Zieky, M.J. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Educational Testing Service.

Machi, L. A., & McEvoy, B. T. (2016). *The literature review: Six steps to success*. Corwin Press.

Menken, K. (2008). *English learners left behind: Standardized testing as language policy*. Multilingual Matters.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241-256. https://doi.org/10.1177/026553229601300302

Mills, G. E., & Gay, L. R. (2016). *Educational research: Competencies for analysis and application*. Pearson Education Limited.

Nicodemus, B., & Swabey, L. (2015). Action research. In Angelelli, C.V. & Baer, B.J., *Researching translation and interpreting* (pp. 157-167). Routledge.

Paltridge, B., & Starfield, S. (2013). *The handbook of English for specific purposes*. Wiley-Blackwell.

Royal Thai Navy. (2017). แผนแม่บทการบริหารทรัพยากรบุคคลกองทัพเรือ พ.ศ. 2560-2579 [Master plan for the development of naval personnel management system for 20-year strategic plan (B.E. 2560–2579)]. Bangkok, Thailand: RTN.

Royal Thai Navy. (2019, April 26). ขออนุมัติปรับปรุงแบบประเมินผลการปฏิบัติงานของกำลังพล ทร. Naval Personnel Department. http://www.person.navy. mi.th/hrm/upload/ ApproveEvaluateNavy62.pdf

Sabieh, C. (2018). English for specific purposes (ESP) testing. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1-6). John Wiley and Sons.

Sahni, S., & Sinha, C. (2016). Systematic literature review on narratives in organizations: Research issues and avenues for future research. *Vision, 368-379*. doi:10.1177/0972262916678085

Shohamy, E. (2007). Language tests as language policy tools. *Assessment in*

    Education Principles Policy and Practice, 14(1), doi:117-130.10.1080/ 09695940701272948

Shohamy, E. (2014). *The power of tests: A critical perspective on the use of language tests*. Routledge. doi:10.4324/9781315837970

Snilstveit, B., Oliver, S., & Vojtkova, M. (2012). Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *Journal of Development Effectiveness, 409-429*. doi: 10.1080/19439342.2012.710641

Willey, T., & Garcia, O. (2016). Language policy and planning in language education: Legacies, consequences, and possibilities. *The Modern Language Journal,* 100(S1), 48-63. https://doi.org/10.1111/modl.12303

JITPICHA JARAYAPUN, PhD is a scholar and executive in the field of Language Teaching development, specializing in language testing for the Royal Thai Navy. Her research interests include language teacher identity, social justice in education, and the implementation of effective language policies.
Email: jitpicha15@gmail.com