ASEAN Journal of Education
Journal homepage: http://aje.research.dusit.ac.th/

A Needs Assessment Study to Improve the Quality of the Thailand National Test

Wisanu Sapsombat*, Arisara Roengsumran

Bureau of Educational Testing, Office of Basic Education Commission, Ministry of Educationes

Article info

Article history:
Received: 17 January 2019
Revised: 28 March 2019
Accepted: 27 May 2019

Keywords:
National Test, Needs Assessment, Standardized Test, Standard of Testing

Abstract

This study uses needs assessment research to assess stakeholders' needs for improving the quality of the Thailand National Test (NT) and its implementation. Related literature was reviewed. To validate the quality of surveys, three experts in the field of educational assessment verified the surveys using questionnaires (Item-Objective Congruence Index) to ensure the content and construct validity. Three types of online surveys with 5-level Likert rating scale questions and open-ended questions were sent to different groups of stakeholders' representatives (staff from the Bureau of Educational Testing (BET) at the Office of Basic Education Commission (OBEC), Educational Service Areas (ESA) and schools). Surveys were conducted to collect opinions about the current state of and expectations about the NT's quality and its implementation. SPSS was used to analyze mean and standard deviation of the results. The modified priority needs index (PNI Modified) formula was utilized to prioritize needs and areas to be improved for enhancing NT effectiveness. Content analysis of results from open-ended questions and a focus group interview indicated factors affecting quality of the NT and strategies for making improvements. Results reveal that of the six domains assessed using the PNI Modified formula, the three domains of the NT with the highest identified improvement needs were the test writers domain (PNI modified = 0.22), the test development domain (PNI modified = 0.16), and the scoring domain (PNI modified = 0.08).

Introduction

Continuous changes in society, science and technology, and economics are inevitable. Some competencies necessary for the workforce today may not have as much significance in the future. We need to evolve our capacity to survive in today's fast-growing world. Education is crucial to prepare productive future citizen. People generally agree that a good education can change someone's life because education strengthens people throung providing knowledge and skills, and influences how they see the world. Therefore, many countries prioritize providing high quality education for their citizens with a goal to prepare them to be ready for and competitive in the future. There are various elements of education systems, and the quality of curriculum, instruction, and assessment, and the link between them, are significant to a successful system.

In Thailand, the Basic Education Core Curriculum

* Corresponding Author
E-mail: 9arisara@gmail.com

was announced in the year 2008, and the curricula in some subject areas such as science and mathematics were revised in the year 2017. The curricula are provided for implementation with appropriate framework and guidance to all educational service area offices, local offices and basic education institutions under jurisdiction of various agencies. The goal is to provide all Thai children and youths with high quality education on important knowledge and skills which will be necessary for students' futures in the constantly changing society and to prepare them for continuous lifelong self-improvement (The Basic Education Core Curriculum B.E. 2551 [A.D. 2008]). Teachers must design their instructions to be consistent with the curriculum to make certain that students will achieve what need to know.

The cycle of teaching and assessment is common. Good assessment requires a variety of measurement tools and techniques such as the classical test with different types of questions from multiple-choices question to essay-type questions, project-based assignments, portfolios, and focus groups discussions. Effective assessments that correspond to curriculum and instructions can give considerable benefits for test takers and test users – mainly students, policy makers, teachers and parents.

There are three functions of educational tests: 1) monitoring at a system level, 2) monitoring at the individual level, and 3) assessment for individual decisions (Wiegers, 2011). Tests are commonly and widely used to evaluate students' learning outcomes. A significant goal of testing is to give useful information for students about their knowledge, skills, and abilities. Well-constructed tests will provide valid and useful data that helps stakeholders to recognize actual students' abilities (how well students know what they are expected to know), students' learning progress, and their needed areas for improvement or for additional support. A famous example is the No Child Left Behind Act of 2001 in the United States, which required that students receive diagnostic reports to allow teachers to address specific diagnostic needs.

In addition, assessment is one vital part of educational quality assurance, as data from assessment reveals how successful the education administration is at accomplishing its goals and objectives. It reflects the success of teachers' instruction and effectiveness of educational administration and management at the school level, the ESA level, and the OBEC level. The information can help policy makers, ESA directors, and school directors to be aware of their strength and weakness, which can help influence better quality decisions about individuals, new policies, new programs, assignment of educational staff, allocation of budgets, teachers' professional development, and other considerations that affect educational quality. In summary, the ultimate goal of educational assessment is to improve educational quality.

However, "not all tests are well developed, nor are all testing practices wise or beneficial, but there is extensive evidence documenting the usefulness of well-constructed, well interpreted tests" (American Education Research Association [AERA], National Council on Measurement in Education [NCME], & American Psychological Association [APA], 2014, p.1). High quality tests provide useful information to improve the quality of education, but inaccurate data from poor tests can cause misunderstanding and misdoing, which can negatively affect the quality of education.

The NT is one of important national standardized tests in Thailand, which is used to test third graders' three domains of competencies (Literacy, Numeracy, and Reasoning Abilities). Educators believe that these domains are students' basic skills and elements for higher thinking skills (Office of Basic Education Commission [OBEC], 2013). Furthermore, they are necessary abilities for students at a young age to be capable of learning advanced knowledge in the future. The Ministry of Education expects that the NT results can assist decision makers (i.e., OBEC, ESA and schools) to make more effective and more relevant decisions for improvement of Thai education quality. However, some stakeholders such as teachers have expressed concerns about the accuracy of NT results and have commented about the quality of the NT in various areas, such as test writers, test administration, and scoring. From literature reviews, most research studies about the NT were conducted to explore strategies to improve students' NT scores and to use NT's findings to promote educational staff like teachers (Apiwongngarm, n.d.; Danaitamonut, 2013). However, there are no research-based studies about the NT's quality issues. As a result, this research study was conducted to explore the NT stakeholders' needs to improve the quality of the NT in Thailand. The study purposes are to identify components of the NT and study the current state of the test both in terms of quality and test implementation. Moreover, the researchers aim to identify stakeholder priority improvement needs for the success of the NT, and to suggest strategies for the improvement of the NT based on the results of the study.

Thailand Educational Assessment

The Basic Education Core Curriculum B.E. 2551 (A.D. 2008) identifies objectives for student assessment to help students to improve their competencies and to assess students' achievement. There are four levels of educational assessments in Thailand: 1) at the classroom level conducted continuously by teachers, 2) at the school level conducted at the end of the semester of and/or the end of the school year to evaluate the success of education services by measuring if students achieve specified goals and to explore areas that need improvement, 3) at the ESA and/or local levels to monitor students under their responsibility by collecting information obtained from schools and from results of the national standardized test provided by OBEC, 4) at the national level, to assess students' achievements at Grades 3, 6, 9 and 12 to obtain information used for comparing quality of education considering different factors (e.g., geographical locations and school sizes). The results help in planning and establishing policies to improve the quality of education.

There are two main organizations responsible to develop national standardized tests. They are the National Institute of Educational Testing Service (Public Organization) (NIETS) and the Bureau of Educational Testing (BET) at the Office of the Basic Education Commission (OBEC) under the supervision of Thai Ministry of Education. The NIETS has developed various tests, such as the Vocational National Educational Test (V-NET) and the General Aptitude Test/Professional and Academic Aptitude Test (GAT/PAT). The most high-stakes tests are the Ordinary National Educational Test (O-NET) developed by the NIETS which are taken annually for 5 subjects (Thai language, Mathematics, Science, Social studies, Religion and Culture, Foreign languages). More than 1,800,000 students took this exam (about 724,000 students at Grade 6, 637,000 students at Grade 9, and 379,000 students at Grade 12) Nation Institute of Education Testing Service. The objective of this exam is to test the knowledge, cognitive ability and academic proficiency of Grade 6, 9 and 12 students according to the Basic Education Core Curriculum B.E 2551 (A.D. 2008). Two other goals are to provide information to the schools to improve their teaching and learning activities and to evaluate the quality of education at the national level.

The Office of the Basic Education Commission is responsible to provide the tests. Well-known tests include a test for reading for first graders, standardized tests for second, fourth, fifth, thirteenth and fourteenth graders, and the National Test (NT) for third graders.

National Test (NT)

The NT is a national standardized test used to test students' competencies at Grade 3, with an objective to assess whether students have acquired adequate essential fundamental abilities as appropriate for their age after completing the first phase of their education to maximize the likelihood of future successes at higher grade levels. The structure of the Basic Education Core Curriculum B.E. 2544 (A.D. 2001) has four phases: Phase one is first to third grades, Phase two is fourth to sixth grades, Phase three is seventh to ninth grades, and Phase four is tenth to twelfth grades (Ministry of Education, Thailand, 2001).

It was first developed and implemented in the year 2012 by OBEC. OBEC has been responsible to develop the test and administer it (deliver it to the test headquarter centers to be picked up by each exam location, analyze results, report results and gave recommendations) since its origin. It is a requirement for students at Grade 3 in schools under the supervision of OBEC and is an option for schools under jurisdiction of various agencies. Most Thai students at Grade 3 around the country (682,446 students) participated in this exam in the year 2017. Students must take the exam at the same time around the country.

The objective of the NT is to assess students' competencies at the national level. The NT tests students in three areas (Literacy, Numeracy, and Reasoning Abilities), which are believed to be substantial fundamental competencies that can assist students to acquire higher knowledge now and in the future. Results of the NT can give various benefits. NT findings reveal the quality of students' learnings, both overall and individually, and reflect the effectiveness of educational administration at the Ministry of Education, ESA, and school levels. The results can be used for decision makings regarding educational plans, policies, strategies, and practices at the national and all other levels. In particular, schools are able to apply each student's individual scores to design instructions appropriate to each student.

The NT test blueprint is designed and developed by experts to align with the current curriculum's learning standards and indicators. One NT comprises thirty questions (27 objective test items and 3 subjective items) with 60-90 minutes test time duration to complete the test.

For the system of assigning scores, computers are used to process the answer sheets. Answers of the subjective test items are scored by local qualified scorers either at the examination locations or at other assigned locations that the test headquarters centers select. The

scorers assign and insert scores for the subjective test items into the same answer sheet that already has students' answers on objective test items. Next, the completed answer sheets are delivered back to the OBEC for the computerized scoring process. The human scoring process for subjective test items must be done during the time period that OBEC specifies. After the exam scoring is finished by the OBEC the NT and the description of correct answers are given to schools. Every year OBEC also provides an annual report including information about the NT results. The report delivers pertinent data for comparing educational quality against different demographic factors (e.g., geographic location, gender, and school size), information regarding strengths and areas which need improvement, and suggestions for improvement. This report is beneficial for planning and promoting the quality of education provided

The NT's results are one important indicator to reflect the effectiveness of education provided to students. The NT's findings have been used widely at all levels to improve the quality of education for students. Invalid NT findings can lead to incorrect decisions regarding policies, resource allocation, teacher development, and so on. Therefore, the NT and its application should be of highest quality, should accurately measure the necessary competencies, and should be fair to all test takers. This study aims to identify components of the NT and study the current state of the test both in terms of quality and test implementation. In addition, researchers want to identify stakeholder priority improvement needs for the success of the NT, and to provide strategies for the improvement of the NT based on the findings of the study.

Needs Assessment

Needs assessment research is a systematic procedure aim at to improving operations, programs, services, performance, and organizations (Altschuld & Kumar, 2010). Needs assessment is a process that attempts to estimate deficiencies, which are measurable gaps between "what is" or the current state of things or situations of interest and the "what should be" or the desired state of things or the situation of interest (Watkins, Meiers & Visser, 2012; Witkin & Altschuld, 1995). It is an effort to identify needs, activities to "gages gaps and insufficiencies", and a method used to predict deficiencies (Royse, Staton-Tindal, Badger, & Webster, 2009). The discrepancy between the current state (what is) and the desired state (what should be) is defined as "need".

Purposes and benefits of needs assessment include clarifying the current state and problems, identifying needs, specifying causes of problems and feasible solutions to them and identifying appropriate intervention and new projects; thus, it can help to create more appropriate policy and programs, improve services and products, and establish or strengthen partnerships (Altschuld & Kumar, 2010; Royse et al., 2009)

Altschuld and Eastmond (2010) proposed three phases of the needs assessment model. They are pre-assessment, assessment, and post-assessment. Each phase includes steps with details to follow (p. 3). Wongwanich (2015) analyzed and synthesized various needs assessment models. The results indicated five stages for a needs assessment model: 1) identifying what should be, 2) exploring the current state of the situation (what is), 3) analyzing discrepancies between 1 and 2, 4) analyzing causes of discrepancies and prioritizing them, and 5) exploring and identifying solutions for the causes identified in Stage 4.

The scope of needs assessment studies are varied. A needs assessment study may conduct only need identification activity. Some may explore the causes of needs (needs analysis) and solutions for the causes of needs (needs solutions). Needs analysis is done to determine the reasons and causes of the "needs" and decide feasible and application solutions (Watkins & Kaufman, 1996).

Needs assessment research has three major stages: 1) identify research objectives for the needs assessment, 2) identify research questions and its limitations, and 3) identify the research framework (Wongwanich, 2015). The research framework contains identification of relevant participants, specification of data collection method and tools, use of data analysis methods, development of reports, and use of results. Needs assessment research can reveal what is happening (current situation), indicate what should be, identify gaps, causes of problems, and provide choices of solutions (Altschuld & Kumar, 2010; Wongwanich, 2015).

Effective assessment can benefit decision makers and practitioners in many ways. Mainly it can improve the quality of decisions made, consequently increasing the probability of achieving desirable goals. For example, it enables them to establish appropriate policies and plan with proper and achievable goals. In addition, the results increase their ability to create effective strategic plans to resolve problems correctly and efficiently, utilizing new solutions that correspond to the current situation and needs of the stakeholders. Accordingly, needs assessment research significantly affects the use of resources more

efficiently.

This study is a needs assessment research. The main purpose is to assess improvement needs for the NT system of Thailand. The research findings can provide useful information for supporting decision making for NT quality improvement. The study was conducted with five purposes: 1) to identify components (elements) of the NT, 2) to study the current state of the NT evaluating test quality and test implementation, 3) to identify the stakeholders' needs and priority improvement areas for the success of the NT, 4) to study factors affecting quality of NT, and 5) to identify strategies for improving the NT.

Methodology

Literature review, surveys and a focus group interview were used to design the study and collect data. Relevant literature was reviewed and synthesized. Publications were reviewed relating to, for example, the NT's background and the current state of educational testing standards established by professionals inside and outside Thailand. Knowledge received helped the researchers to understand the NT and its situation, and to frame this research study.

For the research framework, researchers analyzed and synthesized standards of testing proposed by four sources to develop a framework for this study. The four sources were: 1) The Standard of Testing by the NIETS (NIETS, 2018b), 2) Development of Systems and Criteria for Accreditation of Systems, Methods, and Instruments of Testing Organizations by Sujiva and colleagues (Sujiva, Pasiphol, Tangdhanakanond, & Panjamawat, 2016), 3) Standards for Educational and Psychological Testing (2014), a joint work of the American Education Research Association (AERA), the National Council on Measurement in Education (NCME), and the American Psychological Association (APA), and 4) Standards for Quality and Fairness by Educational Testing Service (ETS) (2014). The research framework contained six domains directly affecting the quality of the NT and its administration. They were: 1) test writers, 2) test development, 3) test administration, 4) test printing and security, 5) scoring, and 6) score reporting and utilization of results.

To assure the quality of the surveys, the researchers provided drafts of the surveys to experts in the field of educational assessment for their review of content, questions, language, layout and format to ensure correctness, appropriateness, completeness and clear understanding. After revising the surveys based on their feedback, questionnaires with the Index of Item-Objective Congruence (IOC) were sent to three experts to verify the content and construct validity. The three experts have doctoral degrees and have experience in the field of educational assessment. The final version of online surveys incorporated recommendations by the experts. The items that had scores lower than 0.8 were revised and the items that had scores higher than or equal to 0.8 were retained (Kanjanavasi, 2005).

After developing and conducting content validity of online surveys, they were used to collect key stakeholders' opinions about the current condition of the NT (test quality and its implementation), expectations for success of the NT, problems with the NT, and factors affecting success of the NT. Three different types of surveys were used to collect staff's opinions for each of three different groups of respondents (staff from the BET, the ESA, and schools). A focus group with twelve participants was employed to obtain more detailed insight into the current state of the NT, including its strengths and weakness, and to acquire suggestions on how to improve the quality of the test and its administration.

Measurement and data collection design

For sample selection, there were 3,691 samples that responded to the surveys. They were: 1) 31 staff from the BET, 2) 366 staff from 183 ESA (which included ESA directors, deputy directors, heads of Division of Educational Supervision Monitoring and Evaluation, and education supervisors responsible for monitoring students' learning achievement), and 3) 3,294 staff from 1,098 schools (which included school directors and teachers (including 3rd grade teachers). Twelve samples were purposely selected to participate in a two-hour focus group interview. They were: 1) OBEC staff at the administrative level, 2) administrators at ESA, and 3) educational supervisors at ESA, 4) school directors, and 5) teachers. Purposeful sampling was employed to select samples for surveys and for the focus group interview.

Phase I: Data was collected using three versions of online surveys for three different groups of respondents according to three different levels of educational administrative structures (OBEC, ESA, and schools). Each group of respondents obtained questions only relevant to their responsibilities (see Table 1). All three different kinds of surveys included both demographic questions and questions with 5-level Likert rating scales asking the respondents' opinions toward the current

situation and their expectations of the NT within six domains. The six domains were: 1) test writers, 2) test development, 3) test administration, 4) test printing and security, 5) scoring, and 6) score reporting and utilization of results. The 5-level Likert scale questions asking respondents' opinions about the current state and their expectations of the NT within six domains were labelled as follows: 5 = Very good, 4 = Good, 3 = Fair, 2 = Poor, 1 = Very poor. All three kinds of surveys contained three open-ended questions asking respondents' opinions about the weaknesses and problems of the NT, factors that may support and influence the success of the NT, and suggestions to improve the NT.

Phase II: A two-hour focus group interview was implemented to gather opinions from twelve participants about the current state of the NT, its problems and obstacles, and how to improve the quality of the NT and its implementation against six domains mentioned earlier.

Table 1: Data collection plan

Questions	Domains of Data Collected	Samples		
		OBEC	ESA	Schools
1. Current state of the NT	1. Test writers	x	-	-
2. Expectations for the NT	2. Test development	x	-	-
3. Open-ended questions (problems of the NT, factors affecting the NT's success, suggestions to improve the NT)	3. Test administration	x	x	x
	4. Test printing and security	x	-	-
	5. Scoring	x	x	-
	6. Reports and utilization of results	x	x	x

Analytical Design

In Phase I, quantitative data was analyzed using descriptive statistics (frequency, percentages, means, and standard deviations) and ANOVA. A Modified Priority Needs Index (PNI) analysis was used for prioritizing needs. PNI modified is defined as (Wongwanich, 2015):

$$PNI = \frac{I-D}{D}$$

I is Importance = mean of expectations
D is Degree of Success = mean of current state

PNI levels (gaps between current state and expectations) of the main six domains and the items within each domain were ranked from the highest to the lowest, which indicated the priority order of stakeholders' needs for improving the NT. The domains or items within a domain that had higher PNI levels should be set as priorities to be improved before those that received lower PNI levels. Content analysis was utilized to analyze qualitative data from open-ended questions. In Phase II, the data collected from a focus group interview was analyzed using content analysis.

Results

Table 2 shows that respondents indicate that the overall current situation of the NT is good ($\overline{X} = 4.29$ S.D. = 0.53). The domain that received the highest average score is the test printing and security ($\overline{X} = 4.45$, S.D. = 0.57). The test administration domain and the scoring domain were at the second ($\overline{X} = 4.35$, S.D. = 0.52) and the third highest average scores ($\overline{X} = 4.29$, S.D. = 0.62) respectively. The two domains that received the lowest scores for the current situation were the test writer domain ($\overline{X} = 3.65$, S.D. = 0.89) and the test development domain ($\overline{X} = 3.93$, S.D. = 0.62). Although they were two domains that were rated lowest, they were still good among respondents' opinions on average. Respondents reported that their overall expectations on the NT is very good ($\overline{X} = 4.54$, S.D. = 0.48). The three domains of the NT that respondents gave highest expectations to be very good were test printing and security ($\overline{X} = 4.74$, S.D. = 0.43), scoring ($\overline{X} = 4.64$, S.D. = 0.53), and test administration ($\overline{X} = 4.57$, S.D. = 0.46). The results showed the three priority domains of the NT with improve needs are test writers (PNI modified = 0.22), test development (PNI modified = 0.16), and scoring (PNI modified = 0.08) (see Table 2).

Table 2. Survey response analysis

Domains	N	Current Situation		Expectation		PNI	Rank
		Mean (D)	S.D. (D)	Mean (I)	S.D. (I)		
Test writers	31	3.65 (6th)	0.89	4.44 (6th)	0.70	0.22	1st
Test development	31	3.93 (5th)	0.62	4.55 (4th)	0.49	0.16	2nd
Scoring	397	4.29 (3rd)	0.62	4.64 (2nd)	0.53	0.08	3rd
Test printing and security	31	4.45 (1st)	0.57	4.74 (1st)	0.43	0.07	4th
Reporting and utilization of results	3,691	4.23 (4th)	0.61	4.50 (5th)	0.54	0.06	5th
Test administration	3,691	4.35 (2nd)	0.52	4.57 (3rd)	0.46	0.05	6th
Total		4.29	0.53	4.54	0.48		

Note. Interpretation of mean scores 4.51-5.00 = very good, 3.51-4.50 = good, 2.51-3.50 = fair, 1.51-2.50 = poor, 1.00-1.50 = very poor

Table 3 shows the highest four identified improvement needs for each of the test writer domain, the test development domain, and the scoring domain. There were a total of nine subareas in the test writer's

domain, thirteen subareas in both the test development scoring domains.

Table 3. Priority needs for improvement of the NT for the test writers, test de velopment, and scoring domains

Test Writers Domain	OBEC Staff (n = 31)			
	D	I	PNI	Rank
Adequate number of test writers.	3.35	4.35	0.30	1st
Test writers are open-minded and willing to improve tests they develop using recommendations from other experts.	3.42	4.42	0.29	2nd
Test writers are knowledgeable and understand curriculum and contents used to construct tests.	3.58	4.42	0.23	3rd
Test writers have teaching experiences and/or have conducted research relevant to the contents used to develop the tests.	3.68	4.52	0.23	3rd
Total (9 subareas)	3.65	4.44	0.22	
Test Development Domain	OBEC Staff (n = 31)			
	D	I	PNI	Rank
Report recommendations and solutions on how to improve tests, select test items, and assemble tests.	3.68	4.58	0.25	1st
Develop documents showing a process of test construction and a manual/guideline for writing test items.	3.65	4.42	0.21	2nd
Assemble tests according to established test structures and test specifications.	3.84	4.55	0.18	3rd
Develop test items according to the item specifications.	3.87	4.52	0.17	4rd
Total (13 subareas)	3.93	4.55	0.16	
Scoring Domain	OBEC and ESA Staff (n = 397)			
	D	I	PNI	Rank
Gathering feedback, especially any issues that occurred, so can solve problems and answers stakeholders.	4.07	4.58	0.12	1st
Develop documents including relevant theories that were used for test and item analysis.	4.11	4.56	0.11	2nd
Develop documents indicating systems for evaluating the quality of the process of checking and scoring answers.	4.15	4.62	0.11	2nd
Create documents describing processes and rationales used in selecting evaluation criteria, interpreting test results, and using evaluation results.	4.13	4.57	0.11	2nd
Total (13 subareas)	4.29	4.64	0.08	

Table 4 shows the items receiving the largest numbers of comments on factors influencing success of the NT, and the items with the largest number of comments on improvement opportunities for the NT, based on responses to the open-ended questions in the surveys.

Table 4. Factors influencing success and improvement opportunities for the NT per open-ended survey questions.

Factors influencing success of the National Test
1. Collaboration among OBEC, ESA and schools.
2. Administrators at all levels are aware of the importance of the NT.
3. Teachers' instructions and support in classrooms aligns with the NT's standards and indicators used to construct the NT.
4. Budgets allocated to organize practice tests for students (mock testing).
5. Administrators encourage and empower teachers to improve students' performance using the NT's results.
6. Staff relevant to the NT are knowledgeable and understand about the NT test and the process of testing administration such as staff at the testing centers.
7. The budget should be sufficient, for example for test administration and results reporting.
8. Parents and students should acknowledge the significance of the NT.
9. There should be payment for scorers who correct answers for subjective test questions at the test centers.
10. Usefulness of the NT Access program.
Improvement opportunities of the National Test
1. There are some mistakes in printing tests.
2. The announcement of the testing date.
3. The process of delivering tests, and the answer keys with descriptions of correct answers (after scoring is complete) is slow.
4. Some pictures in the test are not clear, so students do not know what they are.
5. Some items in the test are too long and not appropriate to Grade 3 students.
6. The test results are announced too late. It should be done before the semester finishes (before students move to the next grade; as a result the utility of results is reduced).
7. The font size used in the test is too small.
8. The level of difficulty for some items is not appropriate for Grade 3 students since students have different levels of performance.
9. The scoring system is not fair and transparent.
10. The budget for printing is too limited.

Respondents from the focus group suggested strategies to improve the NT as follows:

1. Qualification and capacity development of the test writers. There should be a mechanism to establish networks among test writers from every relevant level (i.e., staff from the OBEC, ESA, schools, and teachers in universities), so they are able to connect and share knowledge and resources. Participants pointed out that test writers must be knowledgeable and have experiences in areas relevant for developing tests such as specific contents of knowledge and students' outcome assessment. Organizations that train test writers must make certain that the test writers understand the curriculum and the test blueprint. They should also be provided updated knowledge and techniques to develop test items that measure higher order thinking skills like the PISA test. A data base system including information about test writers at different levels (schools, ESA, national, and international) that is beneficial for test writer's recruitment, selection, and professional development needs to be

created. The data base system can help the OBEC to obtain an array of test writers with different experiences and skills. The OBEC should hire high performance test writers who are accepted at the international level to mentor test writers.

2. Well-constructed tests. The NT should emphasize on situational questions measuring higher order thinking skills like the PISA test. The test writers must develop tests following the standards of test construction and application by identifying and utilizing test purposes, test specifications, item construction, pretests, item evaluation, test assembly, and test administration (Wiegers, 2011). The NT should include more subjective test items like short answers and essays to encourage critical thinking by students. The process of NT development must ensure that tests are not biased and are created following the test blueprint. The test items in the NT must cover the significant learning standards and indicators stated in the current curriculum. The NT should be administered as online computer-based testing for the convenience of the test takers, scoring answers, and reporting test results.

3. Improving the system to check answers. The system to check answers and to process scores should be dependable and safe (secure). The OBEC should have a mechanism to guarantee that the test checkers have/possess the same standards to give scores for answers of the subjective tests items. The results of test analysis should be announced to the public to review its quality.

4. Security of test printing, delivery and control. The test printing, test packing, and test delivery process from central to the test headquarter centers and vice versa should be very highly secure (confidential) with clearly assigned accountable persons for each relevant task.

5. Improving a system to report and use test results. The OBEC should create a system that assists and enables teachers to analyze test results for each student and for all 3rd graders. This will empower teachers to group students according to their abilities. Teachers will be able to plan appropriate and challenging learning activities for each student who are behind and who are gifted. The NT report should provide information that is timely, that is correct, complete and useful, that is easy to access through different channels, and that responds to all levels of stakeholders' needs. There should be a system to offer useful advice regarding how to use the test results to benefit the improvement of students' ability. Handbooks and guidelines on how to interpret and use the test results should be clear and informative.

6. High-standard and transparent test administration. Relevant organizations should provide testing manuals and guidelines that are clear and correct. Research about test implementation should be conducted and the uses of research results should be encouraged among direct stakeholders. There should be more accessible channels to access information about the test such as websites, hotlines, emails, and letters. The NT Access program's performance should be improved (NT Access is a program used to administer the NT including managing test locations, delivering test results, etc.).

Conclusion and Discussion

This research study aims to assess key stakeholders' needs for improving the quality of the NT and its implementation utilizing needs assessment. Related literature was reviewed to frame the research study. Researchers framed questions for data collection against six domains: 1) test writers, 2) test development, 3) scoring answers, 4) test printing and security, 5) reporting and utilization of results, and 6) test administration. Surveys and a focus group interview were conducted with respondents from OBEC, ESA and schools. They were used to explore respondents' opinions about the current state of and their expectations toward NT quality and its implementation as well as to identify needs and priority areas for improvement for enhanced success of the NT. Respondents reported on factors that may influence quality of the NT and recommendations for improving the NT. For the current state of the NT, respondents reported that overall the current state of NT is good and all six domains were rated good, although the test writer domain and the test development domain received the lowest scores. The top three priority domains of the NT for improvement according to the respondents' opinions were test writers (PNI modified = 0.22), test development (PNI modified = 0.16), and scoring (PNI modified = 0.08).

Respondents emphasized that the OBEC should have a mechanism to guarantee that the test writers are highly-qualified because this will significantly affect the quality of the test. Buck (2009) concluded that competent test writers are essential to produce tests with high professional standards. NIETS have organized standardized testing, and computer and paper-based tests, to evaluate teachers' competencies of learning assessment. OBEC could consider using test scores from that test as one criterion for selecting test writers. Additional qualification

requirements should be set to select test writers for a specific test subject such as English language, as recommended by experts that "general proficiency language testers are likely to possess linguistic knowledge, language skills (fluency in the target language, or access to it through a colleague), measurement, and research design skills" (Davies, 1997, p. 82).

Respondents emphasized that the BET staff who are directly involved in developing tests (the OBEC team that is responsible to make tests by recruiting test writers to write test reports) and the test writers should rigorously review and follow the standards of testing that professional organizations and experts have developed. The OBEC must ensure that tests are not biased and are created following the testing blueprints. They stated that it is important that the testing manuals and guidelines are clear and correct. In addition, the NT Access program's performance should be improved, and the OBEC should have a quality control system to check if the test checkers have/possess the same standards to give scores for answers on the same subjective test items.

Participants in the study suggested to establish networks among test writers for exchanging experiences and communication. The OBEC should also develop a data base system including information about test writers to assist the process of test writers' management (recruitment, selection, evaluation, etc.). To increase the quality of the NT, they recommended to include test items that measure higher order thinking skills, such as including more subjective test items like short answers and essays to encourage students to use critical thinking skills. Online computer-based testing for the NT should be created for the convenience of the test takers and test administrators; for example, for scoring answers and reporting test results. Respondents requested more accessible channels to access information about the NT such as testing information, NT results, etc. They also would like to have a system that enables and is user-friendly for teachers to analyze their student's test results, which may link to the learning indicators in the curriculum. This could for example enable teachers to group students according to their abilities in each standard.

The researchers found that there was a statistically significant difference in opinions about areas that need improvement across the six domains among the three groups of respondents (staff from OBEC, ESA, and schools) ($p=0.01$). Looking at individual domains, for example, there was a statistically significant difference in opinions towards the reporting and utilization of results domain among the three groups of respondents ($p=0.02$), but the opinions towards the test administration domain were not statistically significant ($p=0.32$). When comparing between two groups, the researchers found, for example, that ESA staff reported their needs for the reporting and utilization of results domain was higher than the school's staff ($p=0.04$). As a result, focus groups with individual groups of key stakeholders should be conducted in the future to obtain in-depth and different perspectives from each participant group. A study about causes for the differences in opinions among the groups should also be explored; however, this can be costly and labor-intensive.

This study had several limitations. The study used only online surveys and one focus group interview. Most key stakeholder groups were invited to participate in this study (OBEC, ESA, and school teachers), but test writers and students were not included. In addition, only the first step, needs identification, was conducted due to limitation of resources, and needs analysis and needs solutions studies were not explored. Suggestions for utilizing the research's findings and for future studies are proposed next.

Suggestions

From the research findings, NT stakeholders indicated several areas of the NT and its implementation that should be improved. However, this research study only focuses on identifying NT stakeholders' needs for increasing NT quality and prioritizing improvements, while eexperts in the area of needs assessment have suggested that complete needs assessment, including needs identification, needs analysis and needs solutions, should be done to provide comprehensive benefits to the evaluand (Witkin & Altschuld, 1995).

Although some research studies have completed only one or two aspects of a complete needs assessment (i.e., Crişana, Paveleab, & Ghimbuluţc, 2015; Klaharn, 2017), many researchers conducted complete needs assessment studies (i.e., Taraworn, 2015; Vejsuwan, Rupsuwannakul, & Maneelek, 2015). To extend the usefulness of this study, a needs analysis and needs solutions should be executed to identify causes and potential possible solution strategies for the causes. Proposed actions to resolve needs and strategy selection should be done considering the organizational capacity, infrastructure, resources (i.e., budget, capacity of implementers and test writers and time) as well as culture

and other relevant contexts (i.e., political conditions and bureaucratic structure).

Moreover, since test writers and students were not invited to participate in this study and only online surveys and a focus group were used to collect data, future research should include test writers and students and additional types of data collection methods in the study to allow researchers to gain more comprehensive and in-depth information. Examples of additional data collection tools and techniques are: interviews, environment mapping, cultural audits, SWOT analysis, and force field analysis (Altschuld & Eastmond, 2010).

Respondents ranked the test writer domain as the area with improvement needs, with concerns expressed about: 1) sufficient numbers of professional test writers, 2) test writers' knowledge about curriculum contents and educational assessment necessary for constructing tests, and 3) experiences on test writing and teaching.At OBEC, the test writers' tasks include reviewing test blueprints and item specifications, writing and revising items and creating rubrics for scoring. Furthermore, they often work with their partners or in a group to improve test items. Obviously they need not only knowledge and skills of test writing, but also other skills, for example, teamwork, negotiation, conflict resolution, and communication skills. Test writers are a crucial part of developing high quality tests. There should be a more detailed study on the current condition of the NT writers' proficiency, and strategies for capacity building and mechanisms to retain only highly-qualified test writers and remove sub-standard test writers. The study could consider establishing necessary knowledge and skills for test writers, a certification system, a recruitment system, a promotion system, and retention strategies. The study should identify possible constraints and potential issues to maximizing implementation and success.

As part of developing qualifications for test writer certification, future studies should review suggested standards and/or qualifications for test development personnel published from various organizations, for example, 1) Certified Educational Assessor (CEA) by the SACE Board of South Australia (2019); 2) Test Writers' or Assessment Specialist Qualification by the Office of the Civil Service Commission (2010); 3) Standards for Teacher Competence in Educational Assessment of Students established by American Federation of Teachers, National Council on Measurement in Education and National Education Association (1990); and 4) Test Writers' Qualification by National Institute of Educational Testing Service (n.d.). There are many studies about important competencies of test writers and strategies to build their capacity for reference for future research, including: Jiraro, 2013; Hamafyelto, Hamman-Tukur, & Hamafyeltoo, 2015).

The NT development should be at a benchmark high quality/standard since the NT is considered one of the most important national standardized tests widely used in Thailand, and as the results have been used for a variety of reasons such as for promoting teachers and for planning strategies for improving students' abilities. Yet, the test development domain was ranked as the second highest with improvement needs. Future research could invite international and local experts on educational assessment to explore in-depth if the NT meets international and/or local standards and guidelines of testing. These studies could consider improving standards and guidelines utilizing global best practices and focused on purposes of tests and test specifications, item construction, pretesting items, item evaluation, test assembly and development of useful manuals and guidelines for the test. Examples of international standards of test construction are 1) system and criteria for accreditation of systems, methods, and instruments of testing organization (Sujiva, Pasiphol, Tangdhanakanond, & Panjamawat, 2016), 2) standards for educational testing at the national level (NIETS, n.d.), 3) ETS standards for quality and fairness (Educational Testing Service, 2014), 4) standards for educational and psychological testing (AERA, NCME, & APA, 2014).

The researchers believe that the study's findings offer valuable information which can be used for making better decisions about improvement of the NT and its implementation. Domains and their subareas that were ranked with the higher needs should be resolved first. OBEC staff should first establish clear goals, clear strategies and clear guidelines, then create action plans for developing solution strategies, communicating plans, and confirming base support. OBEC staff should communicate the needs assessment results, action plans, and the process to improve the NT quality to relevant stakeholders. After executing plan, they should regularly monitor if improvements follow the established plan, and, finally, evaluate the overall needs assessment efforts (Altschuld & Eastmond, 2010; Wongwanich, 2015).

References

Altschuld, J. W., & Eastmond, J. N. (2010). *Need assessment: Phase I getting started.* Thousand Oaks, CA: SAGE Publications.

Altschuld, J. W., & Kumar, D. D. (2010). *Needs assessment: An overview.* Thousand Oaks, CA: SAGE Publications.

American Education Research Association (AERA), National Council on Measurement in Education (NCME), & American Psychological Association (APA). (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students established.* Retrieved from https://eric.ed.gov/?id=ED323186.

Apiwongngarm, R. (n.d.). *A study of supervision strategies to increase 3rd grade students' National Test literacy scores in the Chiang Rai Educational Service Areas #2.* Retrieved from http://evaluate.cri2.go.th /critwosys/ files/Article_Russada.pdf.

Buck, G. (2009). *Challenges and constraints in language test development.* In J. C. Alderson (Ed.), The politics of language education: Individuals and institutions (pp. 166–184). Bristol, UK: Multilingual Matters.

Crişana, C., Paveleab, A., & Ghimbuluţc, O. (2015). A need assessment on students' career guidance. *procedia-social and behavioral sciences, 180,* 1022-1019.

Danaitamonut, W. (2013). *A study of the appropriateness of using O-NET and NT results for promoting teachers.* Retrieved from http://www.moe.go.th/moe/upload/ news_research/FileUpload/34601-9230.pdf

Davies, A. (1997). The education and training of language testers. *Melbourne Papers in Language Testing, 6*(1), 79–85.

Educational Testing Service (ETS). (2014). *Standards for quality and fairness.* Retrieved from https://www.ets.org/ s/about/pdf/standards.pdf.

Hamafyelto, R.S., Hamman-Tukur, A., & Hamafyelto, S.S. (2015). Assessing teacher competence in test construction and content validity of teacher made examination questions in commerce in Borno State, Nigeria. *Scientific & Academic Publishing, 5*(5), 123-128.

Jiraro, S. (2013). *Development of teacher test item construction competency development model using teacher empowerment through action research process*. (Doctoral dissertation). Retrieved from https://library.car.chula.ac.th/.

Kanjanavasi, S. (2005). *Theories of evaluation.* Bangkok: Chulalongkorn University.

Klaharn, R. (2017). The need assessment for improving the competence of Thai teachers in the measurement and evaluation of analytical thinking. *International Journal of Teaching, Education and Learning, 1*(2). Retrieved from http://grdspublishing.org/index.php/PUPIL/ article/view/958.

Ministry of Education, Thailand. (2001). *The basic education core curriculum B.E. 2544 (A.D. 2001).* Bangkok: Kurusapa Ladprao Publishing.

National Institute of Educational Testing Service (NIETS). (2018a). *Standard of testing.* Retrieved from http://www.niets.or.th

National Institute of Educational Testing Service (NIETS). (2018b). The 2017 *the National Institutes of Educational Testing Service's Annual report.* Retrieved from http://www.niets.or.th/uploads/content_pdf/.

National Institute of Educational Testing Service (NIETS). (n.d.). *Standards for educational testing at national level.* Retrieved from www.niets.or.th.

Office of Basic Education Commission (OBEC). (2013). *Definitions of students' abilities: literacy, numeracy & reasoning abilities.* Retrieved from http://www.jd.ac.th/ news-detail_1122_3142.

Office of the Civil Service Commission. (2010). *Test writers' qualification.* Retrieved from http://person.ddc.moph.go.th/select.pdf.

The SACE Board of South Australia. (2019). *Certified Educational Assessor (CEA).* Retrieved from https://www.sace.sa.edu.au/web/iea/certified-educational-assessors.

Royse, D., Staton-Tindall, M., Badger, K., & Webster, M. (2009). *Needs assessment.* NY: Oxford University Press.

Sujiva, S., Pasiphol, S., Tangdhanakanond, K., & Panjamawat, T. (2016). Development of system and criteria for accreditation of systems, methods, and instruments of testing organization. *SDU Research Journal Humanities and Social Sciences, 12*(3). 183-206.

Taraworn, K. (2015). *A need assessment of the teachers for developing secondary school students' learning assessment.* (Master degree thesis). Retrieved from https://library.car.chula.ac.th/.

Vejsuwan, S., Rupsuwannakul, P., & Maneelek, R. (2015). *Needs assessment in development of quality organization management model in primary educational service area offices.* Retrieved from http://journalgrad.ssru.ac.th/index.php/5-03/article/view/226/160.

Watkins R., Meiers M, & Visser Y. L. (2012). *A guide to assessing needs: Essential tools for collecting information, making decisions, and achieving development results.* Washington DC: The World Bank.

Watkins R., & Kaufman, R. (1996). An update on relating needs assessment and needs analysis. *Performance Improvement, 35*(10), 10-13.

Wiegers, J. (2011). *Eight steps in educational test construction.* (PowerPoint Presentation).Visit to IPST Bangkok on August 18, 2011.

Witkin R., & Altschuld, J. W. (1995). *Planning and conducting needs assessments: A practical guide.* Thousand Oaks, CA: Sage.

Wongwanich, S. (2015). *Needs assessment research* (3rd ed.). Bangkok: Chulalongkorn University Printing.