Developing English for Nurses in Indonesia: From Learning Supporting Tools to Assessment

I Made Rai Jaya Widanta

Mechanical Engineering Department, Bali State Polytechnic, Kampus Bukit Jimbaran Street, Kuta Selatan, Badung-80364, Bali, Indonesia

I Made Ardana Putra

Administration Department, Bali State Polytechnic, Kampus Bukit Jimbaran Street, Kuta Selatan, Badung-80364, Bali, Indonesia

I Wayan Dana Ardika

Civil Engineering Department, Bali State Polytechnic, Kampus Bukit Jimbaran Street, Kuta Selatan, Badung-80364, Bali, Indonesia

Ni Nyoman Ayu Tri Hidayanti

English Letters, Bali Dwipa University, Bali, Indonesia

Abstract—This study aimed to develop English for nurses in Indonesia. Specifically, the development included three levels of a test and their supporting tools such as a competency standard, a scoring rubric, curriculum, and a blueprint. The development was conducted for some stages. The research team was comprised of two lecturers, one staff member, and two students. The students were involved in creating the online version of the test. All supporting tools were validated prior to developing and validating the assessment tool. Upon the tests' validation, the supporting tools were then disseminated to nursing students in Bali province and East Java. The nursing students tried the tests for approximately two hours and 40 minutes. The written test, which consisted of listening, reading, and writing, was done for 2 hours and 30 minutes, and the spoken test was conducted for 10 minutes for a couple of students. Following the try-out, the students were also assigned to fill in the questionnaires distributed via Google Form. A public experiment involving 20 lecturers also gave the lecturers a questionnaire to fill in. Both questionnaires required the respondents to provide opinions, suggestions, and comments on the test before the tests were revised. In conclusion, the tests were valid and submitted to the English Competency Test Certification Foundation under the Directorate for Course and Training Advancement, Republic of Indonesia.

Index Terms—assessment tool, English for nurses, competency test, project-based learning

I. INTRODUCTION

The strategic effort undertaken by the government of Indonesia to struggle for and recover economic condition during and especially upon the COVID-19 pandemic era has been to decrease imports and optimise exports. In order to achieve this goal, the government is empowering industries to produce internationally competitive tangible and intangible products. One of these, namely qualified manpower, is in the service area. Many offers in the form of job vacancies especially in the nursing area are regularly forwarded to Indonesia. Nurses are required to help the elderly, work in orphanages or hospitals, provide home care, or be caregivers. Job providers abroad, both government-owned and private foundations, require that job seekers be competent in English for nursing as evidenced by a certificate of competence.

The presence of this specific requirement has seemed demanding for the job seekers because the government and institutions in other countries including those providing the vacancies have not previously required such English competency tests. Thus, all parties are energised to create a standard test to measure candidates' English competency in the nursing area. The endeavour to construct such an assessment tool has been a focus of attention for both the government and scholars as well as English instructors, educators, and researchers. Nevertheless, the target has not come into reality for reasons including shortages of resources, finances, and information about the needs of stakeholders as well as other related information.

The English Competency Test Certification Foundation under the Directorate for Course and Training Advancement, Republic of Indonesia has been assigned to organise the development of English assessment tools in the country. Of the tools successfully developed, none are considered to be a nursing-related assessment. Ironically, requests for an English competency test in the field of nursing increase in line with the increase in manpower needs in the nursing field in some

other countries. In addition, the need for English courses, training, or materials related to nursing from private language foundations, English courses, training centres, and the like have also increased from year to year.

The development of an English competency test for nurses has been undertaken in a limited and partial manner in the last two decades. It might be caused by its specific use, people restricted request upon it, and its unfamiliarity among its users. Such a test has neither been found and issued to be used formally in any institution nor publicly sold. Although there have been some academic efforts like reviewing, validating, examining, and even developing nursing-related tests, they have been only partially undertaken. Some studies have been conducted on developing English for nurses' learning materials (Putra et al., 2019), analysing skills learners need in English proficiency tests (Yuyun et al., 2014), developing English speaking materials for nursing students (Fadillah et al., 2018), developing English for nursing materials based on communicative language teaching (Febrijanto et al., 2017), and developing and validating English language teacher competence tests (Pishghadam et al., 2011). However, a wholistic effort to develop an English for nursing competency test and its supporting tools which can help institutions, teachers, instructors, and lecturers instruct, assess, and develop curriculum and materials for English for nurses courses has not yet been made.

Some mere obstacles explained in prior seem to be strongly supportive of the development of such an assessment tool. The Indonesian government relied on the English Competency Test Foundation to develop such a test in order for them to have a tool of measurement for the nurses planning to apply for jobs in other countries. These nurses shall be prepared and provided with training or courses in English for nurses. However, the assessment tool for measuring such competency had not been provided. The government wishes that English courses and trainings would be undertaken not only by government-owned foundations but also by private institutions; however, a number of essential components have not yet been provided and determined, such as learning materials, graduate competence, curriculum, assessment rubrics, and assessment tools. In addition, some research undertaken to support the teaching of English for nurses explained in advance was partially and limitedly done. Some of those studies focused on learning materials in schools and some on analysing learners' or test takers' needs and competency, developing speaking materials, developing English nursing materials based on communicative language skills, and validating English language teacher competence tests. This lacuna drove us to undertake a development of a holistic English for nursing assessment comprising test-supporting components such as curriculum, a graduates' competence frame, assessment rubrics, and assessment tools.

II. RELATED STUDIES

A. The English Competency Test and Its Supporting Components

There has been vagueness of comprehension of the difference between proficiency tests and competency tests among people. Basically, an English proficiency test is that which is used to measure one's level of English mastery. This test is used to reveal students' English proficiency or mastery, which is indicated with scores converted into different levels. English tests such as TOEFL, TOEIC, IELTS, and TOEP are of this kind. In contrast, an English competency test is a tool which measures one's English competency. A competency test is normally used to measure the depth of someone's understanding of the given topics in a course. English competency for non-English major students is the ability to use any strategies effectively and comprehensively by utilising reading, writing, listening, and speaking as well as translating using academic, global, linguistic, and intercultural knowledge in nay academic activities in class or at work (Zhang, 2021). Thus, test takers should be given training in line with the topics being evaluated in the form of a test. The test result is not shown in the form of numerical scores but rather in the form of an evaluation of 'competent' or 'not competent'. Thus, a competency test requires a descriptor for the examiner to determine the score of test takers.
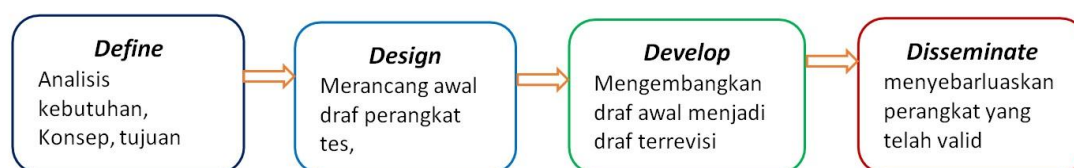
According to Zhang (2021), in developing English competency for a specific field, it is necessary to include aspects of content or subject matter in the test. The content shall be inserted in test-supporting tools such as a graduates' competency standard, descriptors, an assessment rubric, and curriculum. Graduates' competency shall be determined because it will provide guidance about what competence the test taker shall achieve. A descriptor is usually attached in the assessment rubric. The descriptor contains information related to the description of each score, namely what the test taker shall accomplish in order to achieve a certain score. There are some perceptions of what the descriptor shall include, such as that it shall include skills and elements (Lado, 1961); ignore context (Chomsky, 1969); and use comprehensive tests (Oller, 1976) and knowledge about linguistics, topics, character, strategy, specific situations (Bachman & Palmer, 2010), and communicative language ability (Yan, 2012). Of all concepts, the concept proposed by Lado (1961) is supportive because it covers all language skills. The assessment rubric has to cover three main aspects, namely criteria, the description, and the scoring strategy (Jonsson & Panadero, 2016). Additionally, Chowdhury (2018) states that a rubric is essential for examiners when evaluating students' performance. It can also help examiners comprehend the goal to be achieved, provide informative feedback, and support them in being consistent and fair as well as enhancing better learning.

B. Development

This development comprised four main tools supportive of learning activity, namely a students' competence frame, scoring rubrics, curriculum, and a test. The model of development implemented that of 4-D, (i.e., defining, designing, developing, and disseminating; Thiagarajan et al., 1974). The model was chosen because it offers clear stages (Bano, 2018) and systematic steps (Arywiantari et al., 2015). In its implementation, the model was combined with project-

based learning which empowered students majoring in information technology (IT) in the role of research team member to support them learning and applying the knowledge they are gaining in university. The development of learning tools using 4-D has been undertaken by many scholars in numerous fields such as mathematics (Nurmanita et al., 2019) and chemistry (Andromeda et al., 2018; Sakdimah & Dewata, 2018) and in 21st-century learning (Djamahar et al., 2018).

The 4-D model seems to be popular among researchers in pedagogical intervention. Some scholars used full stages, and some adopted part of the whole. Some applied it for English learning materials and some others focused on areas apart from English. Some other researchers utilised 4-D combined with other learning models. Nurmanita et al. (2019) developed a learning model of mathematics which involved project-based learning that was found successful. However, learning mathematics is different from learning language, which involves productive and receptive skills. Sakdimah and Dewata (2018) developed a module for chemistry pursuant to 4-D, and Andromeda, Lufri et al. (2018) developed chemistry materials using 4-D paired with the guided inquiry model. Both experiments produced visible learning materials and models. None of this research was in English language learning, which was apart from English both in model and learning tools. This development of learning tools and a test of English for nurses adopted 4-D by Thiagarajan et al. (1974) with the following stages:

Define
Analisis kebutuhan, Konsep, tujuan → **Design** Merancang awal draf perangkat tes, → **Develop** Mengembangkan draf awal menjadi draf terrevisi → **Disseminate** menyebarluaskan perangkat yang telah valid

C. Project-Based Learning

Project-based learning (PjBL) is an innovative learning approach which provides students with a problem to solve in a project (Thomas & Mergendoller, 2000; Larmer et al., 2003; Bell, 2010). This approach utilises students' prior knowledge so that students are able to discover a concept on their own (Weegar & Pacis, 2012), facilitates students working out their task in a concrete activity (Kriwas, 2007), uses authentic materials in the real world (Ndon, 2011), and prioritises their learning experiences (Kotti, 2008). Through activities such as exploration, discussion, writing reports, and making presentations (Gu Ven, 2014), students are able to activate their communicative skill (Harun, 2006) and improve their autonomous learning skill (Imtiaz & Asif, 2012).

PjBL has been proven effective to improve students' English competency (Imtiaz & Asif, 2012; Rochmahwati, 2016). Apart from its function to improve students' cognitive skill, PjBL has also succeeded in improving their supporting skills such as enthusiasm, self-confidence, creativity, collaboration (Astawa et al., 2017), emotional quotient, learning ability, and enjoyment in learning (Fragoulis, 2009) as well as students' autonomous learning and integrated curriculum development (McCarthy, 2010; Lam, 2011), students' critical thinking (Allan & Stoller, 2005), and vocabulary mastery (Shafaei & Rahim, 2015). Doing projects enables learners to use prior knowledge and energises their language acquisition through group work activities such as discussion, work-presentation inform of the group and other groups. The activities which could attract students' active participation in using their schemata to build new knowledge make the learners pragmatically competent, meaning that they are able to use the language appropriately and effectively (Widanta et al., 2018). In line with this, the success of learners' language acquisition also depends on how teachers give inputs; thus, teachers must provide comprehensible inputs. Inputs benefit learners when they are explicit and comprehensible by means of problem clarification, explicit concept building, and learners' awareness building (Widanta et al., 2020). The more comprehensible the inputs are, the more easily they are processed in the learners' brain and made intakes which will enhance good output. In this case, teachers really have to make learners able to notice as efficiently as possible (Schmidt, 1990; Widanta, 2017).

Students were involved in this research to help develop the online version of the test. The IT students were assigned to participate in developing an online test using the developed offline test. The assignment was in the form of a project done in cooperation with lecturers under the research team including preparing for test development, searching for and collecting materials for the test, and designing the test. The project was also used as the students' final project as a requirement for graduation.

III. DEVELOPMENT AND DISCUSSION

A. Development

The test development covered some stages because there are a number of test-supporting tools which should be produced, namely doing needs analysis; developing a standard of competency, scoring rubric, and competency-based curriculum; developing a test; validating the standard of competency; trying out the test; and conducting a public test.

(a). Needs Analysis

The initiating stage which was the starting point of all test-developing stages was the needs analysis (NA). NA was done in an attempt to recognise the needs of test users. Thus, in undertaking NA, a number of users including main users and sub-users were involved to provide information about their needs regarding the test. The main user of the test is the English Competency Test Foundation. As the research partner to whom the research output will be granted, the foundation reserves a right from the Directorate for Course and Training Advancement, Republic of Indonesia to utilise the assessment tool to assess the English competency of nurses across the country. The foundation successfully gave the research team general and specific guidance for developing the test. With the assistance of the Directorate for Course and Training Advancement, Republic of Indonesia, the foundation encompassed the test development team namely general issues, format, and general content related to language aspect.

Other sub-users which significantly contributed are parties who offered content input for the competency standard, curriculum, scoring rubric, and test. These parties were lectures at a nursing academy; the National of Indonesian nurse board; teachers from a vocational nursing high school; and nurses working in hospitals who directly interact with the academic side such as teaching, learning, curriculum, and practicum, namely nurses working in hospitals and clinics. The input about content was significantly supportive to embody an appropriate standard of competency of nurses as they introduce the current work of nurses, which is apart from that of doctors or midwives. Apart from a direct method of raising inputs, the research team also searched for content from the syllabi, lesson plans, and learning materials of a number of universities, diplomas, and schools offering nursing programs. The NA succeeded in formulising the content of the competency standard for nurses, the scoring rubric, the curriculum, and the test.

(b). Competency Standard for Graduates

The competency standard for graduates was the first supporting tool developed. This was the basis for formulising other test components. Because the test was made to be integrative and interdependent with other tools such as the learning materials, curriculum, scoring rubric, and competency standard for graduates, the test relies on those items.

There are three levels of the standard of competency pursuant to the levels of the test. Each competency standard has some aspects such as competency unit, competency element, and passing indicator, and the passing indicator covers three subcomponents, namely work competency, knowledge, and attitude.

The competence standard formula was gradually developed with some revision sessions. First, it was developed through some focused group discussion sessions resulting in the competency element containing some work or duties held by nurses. Upon some input from stakeholders, the final formula was successfully determined. The competency units finally contained four language skills, namely listening, reading, writing, and speaking because the test focuses on assessing participants' language, not on content or core competency.

The competency element was then agreed to contain subskills in language that the test takers have to gain, such as expressing or explaining the main idea, explaining the meaning of words or phrases in a certain context, writing words or phrases with correct spelling, writing discourses, using grammar and structure appropriately, using vocabularies appropriately, and pronouncing words or phrases correctly.

The passing indicator contains work or duties the nurses have to complete when using their language skills. It is divided into three parts because there are three levels. Eight items of nurses' work or duties are included in the level 1 standard of competency including (1) starting dialog and exchanging personal and information about a patient; (2) asking and giving information about patient condition; (3) explaining the use of medicine, blood pressure, temperature, and patient height and weight; (4) asking about the foods and drinks consumed by the patient; (5) explaining the patient's pattern of activity and exercises; (6) explaining health facilities in the hospital; (7) reporting patient condition to a colleague and/or doctor; and (8) asking about and reporting on the patient's house, family, and environment condition. The level 2 standard of competency covers eight items of work nurses, namely (1) holding further dialog with and about a patient; (2) conducting dialogue via phone with the patient, patient family, and doctor or colleague; (3) explaining orientation and rules about specific rooms; (4) asking about and handling complaints; (5) planning a nursing treatment; (6) explaining the result of a nursing medical record; (7) responding to patient or patient family requests; and (8) giving instructions during a check-up or nursing treatment. Finally, the level 3 competency standard consists of seven items of work including (1) asking about the condition of a patient with a specific diagnosis; (2) offering services or treatment for a patient at a specific room such as the intensive care unit, high intensive care unit, or medical surgery room; (3) giving more complex instructions to a patient in a specific room such as the intensive care unit, high intensive care unit, or medical surgery room; (4) interviewing and observing patients with a specific disease diagnosis; (5) describing the health of patients with a specific disease diagnosis; (6) explaining treatment given by nurses to nurses or doctors; and (7) reporting the result of nurse treatment given to a patient with a specific diagnosis.

The second subsection of knowledge in the language area covers some items that the test takers need to be accurate or appropriate in such as using vocabulary, terms, discourses, content, and knowledge of sentence structures and making summaries or conclusions. Finally, the last subsection, which other standard English tests do not include, is attitude. Attitude is also realised with indicators, namely accuracy, appropriateness, and discipline in performing the four language skills (listening, reading, writing, and speaking) when performing nursing work.

(c). Scoring Rubric

The scoring rubric is essential because it provides a clear answer about what scores or values should be given to a certain test taker's work. The finally agreed-upon scoring rubric uses the polytomy scoring system, which provides three score levels of 0, 1, and 2. Sore 0 represents that less than 65% of the competency required is achieved, score 1 means that the test taker performed 65 to 79.9% competency, and score 2 means that the test taker performed 80 to 100% competency. This system is wiser and more reasonable than that of the former rubric (i.e., the rubric prior to revision), which used the dichotomy system, applying only two scores of 0 and 1. This system suggested an absolute model which categorised test takers' work into two categories: correct and incorrect. The polytomy scoring system is considered more flexible in that each work of the participant represents a certain competency and is valued.

The format of the scoring rubric also contains aspects like competency unit, competency element, passing indicator, scoring component, and quality of score (i.e., 0, 1, or 3). Examiners will use this rubric to score speaking and writing test results. Because the scoring is done using Excel, this scoring system is inserted computer-based scoring program. Thus, the examiner has to input the scores given to each part, and the value for each part as well as the total value in the form of numbers will automatically appear.

(d). Competency-Based Curriculum

The curriculum developed in this study is intended to help with the English learning facilitated by English courses, foundations, vocational schools, and higher nursing education. It will be very helpful for instructors, teachers, lectures, and other professionals in developing learning materials and undertaking the teaching. This curriculum was developed to help learners achieve competency and is thus called a competency-based curriculum.

The curriculum covers some aspects such as the graduate profile; learning outcome; and curriculum matrix, which consists of the competency unit, competency element, materials, credit, duration, method, passing indicator, and module for references. The graduate profile is divided into three parts based on the three levels. Graduate profile level 1 is equalised with level B1 in the Common European Framework of Reference (CEFR), which means having attitude, ability factual knowledge and concept in communicating and doing social and professional work using simple sentences, structure, and vocabulary in performing skills of listening, reading, writing, and speaking. Graduate profile level 2 is equalised with level B2 in CEFR, which means having attitude, ability factual knowledge and concept in communicating and doing social and professional work using compound sentences, structure, and vocabulary in performing skills of listening, reading, writing, and speaking. Graduate profile level 3 is equalised with levels C1 and C2 in CEFR, which means having attitude, ability factual knowledge and concept in communicating and doing social and professional work using compound sentences, structure, and vocabulary in performing skills of listening, reading, writing, and speaking as well as being able to give reasons and opinions and make presentations.

The learning outcome comprises three elements, namely attitude and value, on-the-job ability, and knowledge mastered. Elements of attitude and value are similar at each level, meaning that students are required to be able to implement attitude and character when communicating in English. On-the-job ability covers different jobs or duties at every level. Students are required to be able to speak English appropriately in performing their nursing-related work. Each level draws on different nursing-related work or duties, and knowledge mastered relates to what language knowledge students have to master in order to pass the level. The knowledge concerned covers the ability to use grammar, phrases, sentences, vocabularies, paragraphs, language function, and linguistic features. Elements of grammar and language function differ at every level.

The matrix of competency-based curriculum consists of some items such as the competency unit, competency element, materials, credit, duration, learning method, passing indicators, and module. The competency unit at each level differs pursuant to its goal; for instance, level 1 covers expressing fact and information, expressing the meaning of words in a certain context, expressing the main topic, and writing words or phrases based on appropriate. Level 2 contains slightly different items, such as writing discourse subjects to context, using appropriate grammar, using vocabularies in accordance with context, and writing a discourse with correct structure. Finally, level 3 contains more complex items such as expressing discourse based on context; using grammar with appropriate structure; using vocabulary; pronouncing words, phrases, and sentences based on regulation; and speaking fluently.

(e). Test

Three levels of tests were developed. The development occurred approximately two months after other supporting tools were settled. All of the tests are comprised of four sections, namely listening, reading, writing, and speaking. The first three sections evolved into a written test, and the speaking section individually evolved into a spoken test.

Section 1: Listening, which occurs for 25 minutes, consists of five parts. Part A: Answering Questions has 10 questions, Part B: Understanding Dialogs has 10 questions, Part C: Understanding Mini Talk has 10 questions, Part D: Writing Short Answers has 10 questions, and Part E: Filling in the Blanks consists of 10 questions.

Section 2: Reading Comprehension is divided into two parts, namely Part A: Answers to Questions and Part B: Completing Sentences and Writing Answers. Parts A and B of level 1 and level 2 consist of 35 and 50 questions and 25 and 15 questions, respectively. However, section 2 of level 3 only consists of Part A: Answers to Questions with 35 questions. It differs from test levels 1 and 2 because it is of the highest difficulty level. It includes a higher level of language knowledge and more complicated vocabularies, technical terms, and other phrases. This determination also led researchers to formulise Section 3: Writing for levels 1 and 2 differently than for level 3. In this section, Part A:

Sentence Completion and Part B: Paragraph Writing were only fully involved in Section 3: Writing levels 1 and 2 with number of questions 20 and 1, 20 and 1 respectively, while Section 3: Writing level 3 only consists of Part B: Paragraph Writing with one question.

Section 4: Speaking was designed separately from the other three sections. Speaking is considered the exclusive part of the test because it is one of the important determining aspects for test takers to be considered competent or incompetent. The speaking point of value must be at least 75 for the participant to be competent. If the test taker's other scores per section reached 80 but his or her speaking score did not reach 75, he or she is not considered competent. The speaking test consists of three sections. Section A is a question and response. In this case, test takers are given roleplay cards requesting them (as new nurses) to introduce themselves to each other. Section B is picture-card talk, which requires a pair of test takers to explain pictures and cards consisting of direction. The test takers have to speak to their partner pursuant to the card. Section C is roleplay, where they have to play the role based on the card scenarios.

Prior to test execution, I made more samples of test formulised into questions reservoir, collection, or bank. The questions for the real test were then taken from that question collection. During the test development, the research team collaborated with stakeholders, particularly with the English Competency Test Certification Foundation. The team undertook a number of intensive virtual FGD in order to obtain input on stages for developing, trying, and testing the raw test. They also helped formulate the listening section, which included a recording process. Each test was made for a time allotment of approximately two hours.

The test was completed with other supporting tools apart from the scoring rubric such as the answer sheet, computer-based scoring program, and answer key. The answer sheet was designed in such a way that it fits the test both number of section and number of questions in each section of every level. The computer-based scoring program is a simple program in the form of Excel which can help examiners note the points test takers gain. It is only used for noting points or scores of speaking and writing tests. Each point is inserted in the column concerned, and the table automatically adds and indicates the total score a test taker gains. In order for the examiner or other administrators of the test to be able to check the result of the listening and reading section, an answer key is provided. To do so, examiners note the mark or point of each section manually on the test taker's answer sheet.

(f). Validation

Validation was undertaken prior to developing and trying the test. Two validating sessions were conducted. First, prior to test development, some tools were validated such as competency standard, curriculum, scoring rubric, and question collection. These tools are essential to be validated because they provide guidance for some parties, such as for instructors, teachers, and lecturers to conduct training in English for nurses; for institutions like English courses, work training institutions, schools, or higher education institutions for nurses to develop learning materials; and for the research team to develop the test. The first validating session involved both a language expert and a nursing-related content expert. The language expert was assigned to check and comment on both the Indonesian and English language used in the competency standard, curriculum, and scoring rubric, while the content expert was assigned to check whether the nurse content of those tools was relevant and synchronised with that of a real-life situation like on campus or in a hospital. Thus, an expert who is a lecturer, member of the Nurses Association of National of Indonesia, and practitioner was chosen and assigned to validate the content.

The first validation resulted in revision after both expert judges made some comments. The comments were delivered for revision of the competency standard as the starting point for development of other tools. The main revision suggested by the expert judges concerned the unit of competency, which should focus on language, not content. Thus, it was decided that only four units of competency would be used, namely listening, reading, writing, and speaking. The other comment was on the scoring rubric, which should use the polytomy scoring model instead of the dichotomy model. Apart from this, the scoring criteria shall also use percentage ranges, namely less then 65%, 65 to 79.9%, and 80 to 100%. In addition, the description should use the term 'accuracy', such as 'accuracy in understanding, mentioning, identifying, writing answers, expressing meaning, using, choosing, speaking', and others.

The second validation session specifically focused on assessing the tests. Upon their completion, the three levels of test were validated. The validation was concentrated on language used in the test, format, fonts and font size, the direction of each test section, and pictures included in the test. There were minor comments about the test, such as those concerning instruction, font type and size, and spelling. Based on both comments, revision of those tools was then undertaken.

(g). Dissemination

Dissemination was done to publicise the tests and their supporting tools. A number of dissemination sessions were undertaken. The first session was done for nursing students at government institutions. Approximately 100 students were informed virtually about the newly issued test. They were participants who tried the test. Dissemination was done by informing them about the issued test and its supporting items, the test structure, the sample test, how to apply for the test, to what institution the test would be assigned, test duration, certificate recognition, and the history of the test development. The second dissemination was made to a group of lecturers teaching in some nursing schools and colleges in Bali province. Apart from being given an instrument or questionnaire to fill in about the competency standard, they were also informed about the published test. The third session of dissemination was done virtually with a group of

students and lecturers outside Bali. The members were not asked to criticise and comment on the test but only to let them know and ask questions about the test.

(h). Try-Out

The try-out was held upon test dissemination. It was done once involving students who were given information about the test virtually via previous dissemination. These students were chosen because they majored in nursing, which is in line with the test area. They could represent other members of nursing academies or higher vocational institutions of nursing because they have similar curriculum. In addition, they were participants who could be easily accessed and coordinated with.

The test try-out was undertaken virtually using a paper and pencil test model. There were 135 students involved in the try-out who were divided into three groups (levels 1–3) involving 35 students respectively. The test was divided into two sessions: the written test session and the spoken test session. Of the 140 minutes, the written test took 130 minutes and the spoken test 10 minutes.

During the spoken test, which was undertaken in the first session, there were examiners who interviewed, delivered test questions in the form of scenario cards, and scored their production and three staff who arranged turn of test participants. The spoken test took 10 minutes for a couple of participants. The written test allotted 130 minutes was administered in the second session. It comprised three sections, namely listening, reading, and writing. Each group of participants was gathered in one room.

Each examiner checked the test result pursuant to his or her group. The examiners had to check and input the writing and speaking test mark into a system on a computer followed by the listening and reading test mark. The system could directly add the total mark, and with reference to criteria concerned, the examiner could decide which participants were competent or incompetent.

Apart from taking the test, participants were asked to fill in a commentary form distributed via Google Form to provide their perspective about the test in terms of level of difficulty; their comprehension of words, sentences, or each question in the test; listening materials; directions of each section; and terms or lexicon used. The inputs were then collected, explicated, and analysed to undertake test revision.

(i). Public Experiment

Inputs from stakeholders—in this case, lecturers teaching nursing subjects—were also raised. Upon dissemination of the competency standard and scoring rubric of levels 1, 2 and 3, the lecturers were given an instrument based on which inputs on both tools were delivered. The instrument contains 14 questions asking whether the items asked are in accordance with both tools. These questions concern rules or laws which underlined them, purposes of the tools, level of competency, graduate profile, work and title of nurses at each level, learning outcome, language knowledge of each level, field of work or duties of nurses at each level, competency units of each level, indicators of graduates at each level, and relevance between each aspect in both tools.

Of 20 members of the public experiment towards the tools, 75% participants stated that each aspect was in accordance with the tools. The rest (25%) mostly stated that the tools were in accordance and appropriate; however, some inputs in the forms of suggestions or questions were forwarded to the research team. One participant suggested that the laws or rules which underlined the competency standard have to be ordered accordingly. Another participant suggested that a statement about being 'able to express opinion in an interprofessional team' should be added to the speaking competency standard. Concerning the law or rule, a participant also suggested that law of UU number 38, year 2014 about a nursing competency standard should be added. Another participant suggested that any statement of '*hasil terapi*' (i.e., result of therapy) should be replaced with '*tindakan keperawatan*' (i.e., nursing treatment). The inputs were used to finally revise the tools the tool which was given input for revision was competency standard, while the scoring rubric was considered valid.

(j). Involvement of Students

A research team conducted development of the test and its supporting tools. The team included lecturers, staff, and two students. The students majoring in informatic management were mainly assigned to develop the online or computer-based test version. Prior to their final task execution, they were involved in some stages of development including gathering data for needs analysis, collecting resources used for determining and formulating the nurse competency standard, and helping with FGD in arranging work-related topics for the competency standard and developing questions for the test. Upon the three-level tests were successfully developed.

The students were assigned to create the online test version. They were given data to be uploaded. The data consisted of the final test-like to be made into the test to provide test-taking candidates with a practice test. The practice test questions were made pursuant to the blueprint of the test. The blueprint of the test was prepared to synchronise the real test with other supporting tools such as the competency standard, curriculum, and scoring rubric. In addition, it correlated between materials of learning and the evaluation tool. The development stages and involvement of students and staff indicated that it was undertaken pursuant to a project-based model (Thomas & Mergendoller, 2000; Larmer et al., 2003; Bell, 2010; Weegar & Pacis, 2012; Kriwas, 2007; Ndon, 2011; Kotti, 2008). Involvement of students in

lecturers' project is definitely suggested to give students opportunities to implement their knowledge and skills obtained during their study.

B. Discussion

Development of this assessment tool was done holistically. Realising the competency test for nurses was the main goal of the development. However, because this assessment tool is used widely not only for test takers but also for other stakeholders as a barometer of whether the English for nurses teaching and learning is considered successful, it was developed with some supporting tools such as the competency standard, curriculum, and scoring rubrics. These supporting tools were used as a device to bridge what test takers had learned during preparation and what tool was used to evaluate their achievement. Coherence between the two main aspects of what was learned and what was used to evaluate became essential in this case.

According to the result of the literature review, this development is considered novel because none of the findings in the same area were like it. This development could foster a complete set of learning equipment from the beginning to the end. Although learning materials will become the further development goal and have not been embodied, it will be a small part of its continuation because other devices such as curriculum and competency standard have been realised. This development did not only realise learning materials of English for nurses (Putra et al., 2019; Fadillah et al., 2018) or focus on drawing skills learners need in an English proficiency test (Yuyun et al., 2014), developing English for nursing materials based on communicative language teaching (Febrijanto et al., 2017), or developing and validating an English language teacher competence test (Pishghadam et al., 2011) but rather all of the above. Almost all of the research outputs embodied by the abovementioned researchers were covered by this development.

The test as the final goal of the development was considered holistic because the test could cover cognitive, affective, and psychomotor aspects. The three aspects are included in the competency standard in more realistic embodiments such as knowledge, ability to perform work or duties, and attitude shown when using language ability at work. Based on the review result, other English tests could not be realised explicitly how aspect of psychomotor was performed in the test. This test could contain aspects of psychomotor with, that how nurses shall use English with an appropriate attitude. It was evidenced with the used of the words 'appropriately' and 'politely' or other attitude-related adverbs of manner.

IV. CONCLUSION

In conclusion, the test development can be considered holistic because it succeeded in developing not only the test but also its supporting tools, such as the competency standard, assessment rubrics, and competency-based curriculum. The development comprised some stages such as needs analysis and developing the competency standard, curriculum, scoring rubrics, and blueprint. Apart from these, the development also covered a number of actions such as validating and testing. Because the materials for learning have not been developed, further research may focus on the materials or book given by the researcher.

REFERENCES

[1]   Allan, B & Stoller, F. (2005). Maximizing the benefits of project work in foreign language classrooms. *English Teaching Forum (43)* 4. 10-21.

[2]   Andromeda, A., Lufri, Festiyed., Ellizer, E., Iryani, I., Guspatni, G., Fitri, L. (2018). *IOP Conference Series: Materials Science and Engineering*, 335.

[3]   Arywiantari, D., Agung, A.A.G., Sastra, I. D. K. (2015). Pengembangan multimedia Interaktif model 4D pada pembelajaran IPA di SMP Negeri 3 Singaraja. *E journal Edutech Universitas Pendidikan Ganesha Jurusan Teknologi Pendidikan 3*(1), 1-12.

[4]   Astawa, N. L. P. N. S, L. P. Artini, & P. K. Nitiasih. (2017). Project-based learning activities and EFL students' productive skills in English. *Journal of Language Teaching and Research. Vol 8*. No. 6. pp. 114-1155. http://dxdoi.org/10.17507/jltr.0806.16.

[5]   Bachman, L. F., and Palmer, A. S. (2010). *Language testing in practice: designing and developing useful language tests.*

[6]   Bano, V. O. (2018). Pengembangan model pelatihan pengelolaan penilaian autentikguru IPA SMP. *Kelola Jurnal Mnajemen Pendidikan. Vol. 5*, No. 2, Juli-Desember. pp. 139-151. jurnalkelola@gmail.com

[7]   Chomsky, N. (1969). *Aspect of the theory of syntax*. Cambridge: MIT Press. http://doi.org/10.21236/AD0616323. Oxford: Oxford University Press.

[8]   Chowdhury, F. (2018). Application of Rubrics in the Classroom: A Vital Tool for Improvement in Assessment, Feedback and Learning. *International Education Studies, 12*(1), 61. https://doi.org/10.5539/ies.v12n1p61.

[9]   Djamahar, R., Ristanto, R.H., Sartono, N., Ichsan, I. Z., and Muhlisi, A. (2018). *CIRSA: Designing instructional kits to.*

[10] Fadilah, U., Sugirin. (2018). Developing English Speaking Materials for Nursing Students. *Advances in Social Sciences. Education & Humanities Research, vol. 326*. 3rd International Conference on Current Issues in Education (ICCIE 2018) pp. 503-508

[11] Febrijanto & Kurnianjati. (2017). Developing ESP nursing based on communicative language teaching. *Journal of English Teaching and Research Vol. 2*, Number 2. October 2017 pp. 62-87.

[12] Fragoulis, I. (2009). Project-based learning in the teaching of English as a foreign language in Greek Primary School: From theory to practice. *English Language Teaching.* www.CCSenet.org/journal.html. Vol 2. No. 3. September 2009.

[13] Gu Ven, Z. Z. (2014). Project Based Learning: A constructive way toward learner autonomy. *International Journal of Languages' Education and Teaching.* Turkey: Necmettin Erbakan University Department of Linguistics.

[14] Harun, Y. (2006). *Project-based learning handbook "Educating the millennial learners"*. Kuala Lumpur: Educational Technology Division Ministry of Education.

[15] Imtiaz, S. & Asif, S. I. (2012). "I almost learnt to learn": Promoting learner's autonomy through project-based learning in access classrooms. *Language in India 12*. 1, 24-45.

[16] Jonsson, A., & Panadero, E. (2016). *Chapter 7 The Use and Design of Rubrics to Support AfL.* January, 1–14. https://doi.org/10.1007/978-981-10-3045-1

[17] Kotti, D. (2008). Experiential learning form theory to practice. *Adult education, 13*, 35-41.

[18] Kriwas, S. (2007). *Pedagogical science. Basic Theme.* Athens: Gutenberg. (in Greek).

[19] Lado, R. (1961). *Language testing: the construction and use of foreign language tests*. London: Longman.

[20] Larmer, J., Markham, T., & Ravitz, J. (2003). *Project-based learning handbook*. Oakland, CA: Wilsted and Tylor.

[21] Li, T., Zhang, Y. (2021). The construct of English competence and test design for non-English major postgraduates. *English Language Teaching, Vol.14*, no.1: elt.ccsenet.org.

[22] Mc. Carthy, T. (2010). *Integrating Project-based Learning into a Traditional Skill-based Curriculum to Foster Learner Autonomy: Actin Reseach.*

[23] Ndon, U. (2011). *Hybrid-context instructional model: The internet and the classrooms: The way teachers experience it.* Information Age Publishing Inc. The USA. 28th Annual Conference on Distance Teaching & Learning.

[24] Nurmanita, Pargaulan, S., and Sitompul, P. (2019). Development of learning device through problem-based learning model assisted by Geogebra to improve students' critical mathematical thinking ability. *Journal of Mathematical Sciences and Applications, Vol. 7*, no. 1, 1-9. http://pubs.sciepub.com/jmsa/7/1/1. November 2019.

[25] Oller, J. W. (1976). Evidence of a general language proficiency factor: an expectancy grammar. *Die neven sprochen, 75*(2), 165-174.

[26] Pishghadam. R, Baghaei. P. & Ahmadi, S.A. (2011). Development and validation of an English language teacher competence tests using item response theory. *The International Journal of Educational and Psychological Assessment Vol. 8*(2) pp. 54-68. September 2011.

[27] Putra, I G.A.P., Padmadewi, N.N., Budasi, I G. (2019). Developing English Materials or Nursing Students of SMK Kesehatan Bali Medika Denpasar. *Jurnal Pendidikan Bahasa Inggris Indonesia.*

[28] Rochmahwati, P. (2016). Project-based learning to raise students' speaking ability: its effect and implementation (mix-method research in speaking II subject at STAIN Ponorogo) Kodifikasia, 9(1), 199-222. *Iranian Journal of Language Teaching Research 3*(2) pp.83-99. www.urmia.ac.ir/jltr. June 2016.

[29] Sakdimah, Dj, L., and Dewata, I. (2018). *Development of chemistry laboratory guides based on chemoentrepreneurship (Cep) for odd semester in science class second grade*. Proceeding ICESST (International Conference on Education, Social Sciences.

[30] Schmidt, R. (1990). "The role of consciousness in second language learning". *Applied Linguistics. 11*. 129-158.

[31] Shafaei, A. & H. A. Rahim. (2015). Does project-based learning enhance Iranian EFL learners' vocabulary and recall and retention? *Iranian Journal of Language Teaching Research 3*(2) pp. 83-99. July 2015.

[32] Thiagarajan, S., Semmel, D.S., & Semmel, M. I., (1974). Instructional development for training teacher of exceptional children. Blomington: Indiana University. Technology), pp.252-263 http://icesst.fipunp.ac.id. empower 21th century skill. *Educational Process International Journal. Vol. 7*, issue 3. pp. 200-208. www.edupij.com.

[33] Thomas, J. W. & Mergendoller, J. R. (2000). *Managing project-based learning: Principles from the field.* Paper presented at the annual meeting of the American Education Research Association, New Orleans.

[34] Weegar, M. A. & Pacis, D. (2012). A Comparison of two theory of learning-behaviorism and constructivism as applied to face-to-face and online learning. In *Proceedings E-Leader Conference, Manila.* https://www.g-casa.com/conference/manila/ppt/Weegar.pdf (accessed 13/11/2016).

[35] Widanta, et al. (2017). How should pragmatic be taught at vocational school? *International Journal of Language and Linguistics. Vol 4.* 2. www.ijllnet.com. pp. 110-115. Retrieved August 5, 2021.

[36] Widanta, et al. (2018). Interlanguage request modification: a case in vocational college. *Journal of Physics: conference series* 953012095, doi:10-1088/1742-6596/953/1/012095, pp.1-8, retrieved June 8, 2021.

[37] Widanta, et al. (2020). Explicit inputs and concept mastery: a case in Indonesian language for foreigner (ILF) instruction. *Asian ESP Journal, Volume 16,* 1 December 2020, pp.101-126. Retrieved August 5, 2021.

[38] Yan, M. (2012). *A study of genre-based discursive competence of business English: construct definition and test development* (Doctoral dissertation) Retrieved from http://.www.cnki.net/

[39] Yuyun, Meyling, Laksana, Abednego. (2018). A study of English proficiency test among the first-year university students. *Journal of Language and Literature, Vol. 18*. No.1, April 2018.

I Made Rai Jaya Widanta is an associate professor in Linguistics and is currently working as a permanent lecturer at Politeknik Negeri Bali. His research interests include developing learning model, first and second language acquisition, and inter-language pragmatics.

I Made Ardana Putra is a permanent English lecturer at Administration department, Politeknik Negeri Bali. His research interests are developing English learning model, developing agro tourism and cultural tourism.

I Wayan Dana Ardika started his career as an English lecturer in Bali State Polytechnic in 2009. His research interests are both in English education and English literature, including designing learning model, designing engineering dictionary, and pragmatics.

Ni Nyoman Ayu Tri Hidayanti is a lecturer at Bali Dwipa University, Bali, Indonesia. She is assigned to teach both general English and English for specific purposes (ESP).