CHAPTER 4

FINDINGS AND INTERPRETATIONS

This section presents about findings and interpretation of the study during the phase of developing ESP reading assessment which include analyses, design, evaluation and revision.

4.1 Findings

4.1.1 Analyses Phase

The result of each analysis in analyses phase is described as follows.

4.1.1.1 Instructional Analysis

In this phase, first of all, the English curriculum used in the school was analyzed. The materials about the descriptive texts in reading assessment were also analyzed in the students' textbook entitled 'Forward' for the tenth and eleventh grade, and 'Get along with English' for the twelfth grade. Finally, the writer determined that the basic competence taken from the English curriculum for the purpose of the study was 3.3 and 4.3 which were about descriptive text.

In addition, the potential topics (nursing content) to be integrated into the English curriculum were analyzed. It was analyzed from the basic competency and the textbooks for nursing assistant curriculum for vocational high school and based on the Indonesia National Working Competency Standards for Health and Social Health Care dealing with the Indonesia National Qualification Framework (Ministry of Manpower and Transmigration, 2007).

Based on the textbooks analysis, it can be concluded that descriptive text was mostly used in the nursing textbook. Then, from the English textbook, it is found that descriptive reading materials were provided, though, most of the content of descriptive text do not cover the students' expertise competence, nursing (see appendix 6). The reading assessment in the textbook was even nothing. It means that the material (reading passages) of descriptive text about nursing content in the book was still insufficient.

4.1.1.2 Students' Needs Analysis

In this phase, the aim of students' needs analysis was intended to find out information about the needs of students in term of ESP (in general) and reading. The result of needs analysis was summarized in four different themes as Target Situation Analysis (TSA), Present Situation Analysis (PSA), Deficiency Analysis (DA), and Strategy Analysis (SA). To find out the information above, a questionnaire adapted from Petrus (2012) consisting of 21 items to 44 students in three different classes was administered. The result is presented as follows:

In Target Situation Analysis (TSA), most of the students realized that English is important in order to communicate with the foreigner (70.5%), and to support their carrier in the future (65%). Half of them (47.7%) realized that English is important to help to continue further study in their field, and the small number of them thought that it is important to pass the national examination (20%).

Almost all of the students expected that English should make them able to communicate fluently using English in oral communication in the future workforce (88.6%). Most of them expect that they can use grammar correctly and master the vocabulary related to the nursing field (56.8%). Almost half of them expect to be able to communicate in written communication in the future workforce (31.8%).

Almost all of the students wanted to use English to do oral communication with colleagues or clients/patients (88.6%) and to deepen the nursing skill by reading the English texts (63%), small number of them wanted to use English to do written communication both in formal or informal contexts (13.6%).

Almost half of the students agreed that their English proficiency level for carrier should be in the advanced level (47.7%), almost half of the other answered intermediate (34.1%) and beginner (18.2%).

Meanwhile, in relation to the students' present situation analysis (PSA), almost all of the students said that their current proficiency level of English was beginner (88.6%), and small number of them were in the intermediate level (11.4%), and there were no students at the advanced level (0%).

Especially in learning reading, most of the students desired to have texts related to the context of the nursing field (65%), and authentic materials (63.6%). Almost half of the students desired to have texts with picture (47.7%), and the small number of them desired to have text with the list of new vocabulary (25%), the small number of others gave another additional opinion (9.1%).

The length of input text desired by students was 100 – 150 words (36.4%), > 200 words (25%), <100 words (25%), and 150 – 200 words (13.6%).

In addition, most of the students desired the topic for learning input in reading was daily life routines (77.3%), the nursing field (61.4%). Almost half of them desired to have education/school topic (45.5%), and issues/latest news (31.8%). Half of them also gave their other opinions (45.5%).

The next, most of the students preferred activity for learning reading was comprehending a text and translating it into Indonesian language (61.4%), almost half of them preferred reading a text loudly with correct pronunciation and intonation (44.7%), discussing and comprehending a text in a group (38.6%), analyzing the meaning of the new vocabulary from the context (29.5%), and reading a text individually and answering questions related to the text (27.3%). One student gave their other additional opinion (2.3%).

Furthermore, almost all of the students stated that the reading text in English textbook used in the classroom was difficult (77.3%). The small number of them stated easy (22.7%) and very difficult (2.3%). No students stated that it was very easy (0%).

Finally, most of the students stated that the frequency of English teacher in providing the reading text/material and reading assessment with the topic related to the nursing field was sometimes (61.4%). Almost half of them stated it was often (27.3%), and the small number of them stated very often (6.8%), and never (4.5%).

The next, it was related to deficiency analysis (DA), most of the students stated that they felt difficult to learn grammar (75%) and listening (75%). Half of them stated vocabulary (52.3%), and pronunciation (52.3%). Almost half of them

chose speaking (47.7%), writing (34.1%), and reading (27.3%). The small number of them stated others (4.5%) such as translation.

To be specific, students felt difficult with reading skill such as comprehending certain words (63.6%), comprehending the main idea (56.8%), comprehending the detail information (38.6%), making and drawing conclusion (38.6%), making conclusion related to the chronological order of event (31.8%), making conclusion related to cause and effect (18.2%).

In relation to strategic analysis (SA), most of the students preferred matching the words or expressions with correct meaning in the Indonesian language (63.6%) as the activity for learning vocabulary. Almost half of them preferred matching the words or expressions with the pictures (38.6%), predicting the meaning of the new words based on the context (38.6%), completing the incomplete sentence or paragraph with the provided words (34.1%), completing the incomplete sentence or paragraph with students' own words (29.5%), and others (27.3%).

In learning grammar activities, most of the students preferred to writing sentences based on structure learned (63.6%), identifying the incorrect sentence structure (54.5%), evaluating the incorrect sentences structure (52.3%). Only a small number of them preferred other activities (22.7%).

In learning pronunciation activities, almost all of the students preferred repeating the words pronounced by the teacher (86.4%). Most of them preferred pronouncing the words based on the phonetic transcription (59.1%). Almost half of them preferred discussing the correct pronunciation in a small group or with other friends (34.1%). A small number of them preferred other activities (22.7%).

Most of the students preferred group work and whole class work (68.2%), individual work (40.9%), pair work (27.3%), and others (6.8%) as class management for completing the task.

Meanwhile, for learning style preference, they preferred discussing with other friends to solve the problem or to do the tasks (70.5%), writing all information given by the teachers (54.5%), listening to the teachers' explanation only (27.3%), and writing all information given by the teachers (13.6%).

Most of the students preferred if the teachers: give examples about the topic which is learned and give an assignment (75%), give students questions to finish and discuss them later (56.8%), supervise the students' works and help them when students get a problem (36.4%), walk around and give comment to students' works (34.1%), and others opinion (13.6%).

In addition, most of the students stated that integrating the topic/content related to the nursing field in reading text was very important (61.4%), and important (38.6%). No students stated not important (0%), and very unimportant (0%).

The students stated agree (47.7%), very agree (40.9, disagree (11.4%), very disagree (0%) that discussing the content of the nursing field in reading text could ease them to comprehend the text.

4.1.1.3 Students' Functional Reading Level Analysis

This analysis was measured in order to match the students' functional reading level with the readability of the product. Group Assessment in Reading by Warncke and Shipman (1984) (Level 2 – 8) was used to determine the students' functional reading level. The result was described in the table 4.1 as follows.

Table 4.1
The Result of Students' Functional Reading Level Analysis

Instructional Level	**Functional Reading Level**					
	Frustration (Correct number ≤ 5		Instructional (Correct number 6-8)		Independent (Correct number 9-10)	
	N	%	N	%	N	%
2	0	0	26	59.09	18	40.91
3	6	13.64	32	72.73	6	13.64
4	14	31.82	24	**54.55**	6	13.64
5	21	47.73	19	**43.18**	4	9.09
6	37	84.09	6	13.64	1	2.27

7	43	97.73	1	2.27	0	0
8	42	95.45	2	4.55	0	0

As shown in table 4.1, half of the students (54.55%) had instructional level 4 and almost half of the students were in level 5 (43.18%). As adapted from Cooper, et al. (1989), and due to the distribution of students' instructional level which is mostly (half of students) in level 4 and level 5, the students' instructional reading level was determined by considering these two maximum levels which was achieved by the students as their instructional reading level. Since almost half of students were also in level 5, then this becomes the basis in determining the instructional level for reading assessment development. Therefore, the assessment passages to be developed include three levels below and three levels above (Level 2, 3, 4, **5**, 6, 7, 8).

4.1.2 Design Phase

In this phase, the reading assessment was developed. The genre of the original text from two e-books was in form of descriptive text (describing process). Therefore, the reading assessment was developed in form of descriptive texts. The product was also developed by adjusting to the students' functional reading level. There were seven adapted passages. The passages had been adapted by simplifying and modifying. Since the students' instructional reading level was mostly level 4 and level 5, the developed passage covered three levels below and three levels above the students' instructional reading level. The readability of each passage was adjusted and calculated using the Raygor Estimate Graph which is available at http://readabilityformulas.com/free-fry-graph-test.php. Finally, there were 7 developed passages that consist of passage level 2, 3, 4, **5**, 6, 7, 8. It can be seen in table 4.2.

Table 4.2

The Passages of ESP Reading Assessment

Passage	Topic	Readability level	Source
1	Bathing	2	Adapted from Textbook of basic nursing (2012) & Nursing assistant training (2013)
2	Bed Making	3	Adapted from Textbook of basic nursing (2012) & Nursing assistant training (2013)
3	Nail Care	4	Adapted from Textbook of basic nursing (2012) & Nursing assistant training (2013)
4	Elimination	5	Adapted from Textbook of basic nursing (2012) & Nursing assistant training (2013)
5	Perineal Care	6	Adapted from Textbook of basic nursing (2012) & Nursing assistant training (2013)
6	Vital Signs	7	Adapted from Textbook of basic nursing (2012) & Nursing assistant training (2013)
7	Hot and Cold Application (Pain)	8	Adapted from Textbook of basic nursing (2012)

Additionally, the writer also constructed the test item for reading assessment. The test items were in form of multiple choice questions (with four alternative choices) which covered some reading aspects such as details, sequence, main idea, cause and effect, inference, vocabulary, critical reading, and study skill: library. Finally, the total constructed test items were 70 (see table 4.3).

Table 4.3

Specification of ESP Reading Assessment

No	Reading Aspect	Test Item Number
1	Details	2, 3, 11, 23, 30, 34, 42, 51, 53, 62, 63
2	Cause and Effect	1, 12, 17, 21, 32, 43, 55, 61
3	Sequence	4, 13, 15, 22, 36, 37, 44, 52, 64,
4	Main Idea	5, 14, 25, 33, 45, 56, 66,
5	Inference	10, 24, 35, 46, 54, 65,

6	Critical Reading	6, 16, 26, 31, 41, 47, 57, 67,
7	Vocabulary	7, 8, 18, 19, 27, 28, 38, 39, 48, 49, 58, 59, 68, 69,
8	Study Skill: Library	9, 20, 29, 40, 50, 60, 70

4.1.3 Evaluation and Revision Phase

The evaluation and revision of the product were done thoroughly to fulfill the required process as proposed by Tessmer (1993). The results of the evaluation and revision phase are described in the following description.

4.1.3.1 Self Evaluation

In evaluating the developed reading assessment, the writer tried to treat herself as an expert, so the evaluation process did not only a simply look over the product. The self-evaluation covered all of the four aspects such as instructional design (construct), language, media (production), and content (subject matter). In addition, the writer evaluated the product to find out some errors and mistakes, for example misspelling, ungrammatical sentences, and punctuation. The writer also changed the complex sentence or words into simple sentence or words in order to match the readability of the developed reading passages with the students' instructional reading level.

The reading assessment after being evaluated in self evaluation was called as prototype 1. Some examples of self-evaluation were presented in appendix 1.

4.1.3.2 Expert Review

The second phase was the expert review evaluation that consisted of four evaluation aspects such as instructional design (construct), media (production), language, and content (subject matter).

The first evaluation aspect in this phase was ***Instructional Design (Construct).*** To develop the product with good construct validation, it is important to make sure whether the validation sheet has been required certain standardize requirement. Therefore, it was developed by adapting from Tessmer

(1993) and also Waugh & Gronlund (2013). The total statements in ID validation sheet were 25 and it was given to two ID experts.

The expert rating the product consisted of two ID experts. Both experts are qualified in their expertise which was constructing the assessment. The first expert stated that the product was **valid without revision**. Minor revisions were suggested, such as to recheck and revise some grammatical errors in the questions, and to revise the test direction. The example of the revision from the first ID expert was described in appendix 2.

In addition, the second ID expert had some notes to revise the face validity of the product, such as the alternatives choices should be begun from the shortest into the longest one. The others are about the capital letter or the lower case in the alternative choices, the use of full stop (three full stop if the answer required in the middle of sentence, and four full stop if the answer is required at the end of sentence), and then the word classes in the alternative choice should be parallel. The distribution of the answer keys should be in the same total number for four alternatives choices and do not make a certain pattern. The expert also suggested writing the table of specification for all of the questions. In the beginning, more than one indicator for one question was written, then the expert asked to revise it in which one question only has one indicator. The expert also checked whether the question item was really in line with the reading skill measured. She suggested revising some questions in order to be more understandable by the students. Finally, she declared that the product was **valid with revision**. The example of some revised questions was described in appendix 3.

The second evaluation aspect in the evaluation and revision phase was ***Media (Production)***. The questionnaire was adapted from Yazdanmehr and Shoghi (2014). After reviewing the product, he stated that the product was **valid without revision**. The media expert gave score 4, except for statement number 7, 9 and 15, he rated 3. The score rated on those three statements were analyzed, such as which are related to whether the lines of the text are appropriate in length and are not soanning more than half the width of the page, whether the visuals have the reinforcing and elaborative function to help convey the thematic issues of

the texts, and whether the composition of colors is attractive and well-designed. From the rating score on media (production) validation sheet, then the adjustment of the lines of the text and the visual of the product were done appropriately, but for the color composition, it will be considered for the future development.

The third evaluation aspect in the evaluation and revision phase was ***Language***. The expert suggested considering another type of reading skill which was study skill: library. This skill was viewed as one of the important skills needed to be developed by the students. Therefore, 7 questions of this skills were added to the early total number of the question (63 questions), so it becomes into 70 question in total. After reviewing the product, she stated that the product was **valid without revision**. The example of the additional questions based on the language expert's suggestion was described in appendix 4.

The last evaluation aspect in the evaluation and revision phase was ***Content (Subject matter)***. To ensure that the content of reading assessment is relevant, the validity of developed reading assessment content was rated by the expert of subject matter (content). At the beginning of the validation process, the expert was attached all of the things should be reviewed. The validation sheet which is consisted of the aspects to be rated was explained to the expert. During the validation process, the expert gave some constructive suggestion for the better revision of the product, especially in content. The expert advised to accommodate and consider the Indonesia culture for certain topics, such as bathing, elimination, and perineal care. The materials (reading passages) in reading assessment were adapted by simplifying and accommodating the culture in Indonesian context for certain topics. After the validation process has completed, the expert decided that the product was **valid with revision**. Some revision examples of the product from the content expert can be seen in appendix 5.

Table 4.4 gave a summary of the validity of the product based on the result of all of the expert review validation.

Table 4.4

Summary of Validity in Expert Review

No	Aspects	Average Score	Category
1	Construct (Instructional Design)	3.54	Very High Validity
2	Media (Production)	3.84	Very High Validity
3	Language	3.48	Very High Validity
4	Content (Subject Matter)	3.79	Very High Validity
	Total Average	**3.66**	**Very High Validity**

As stated in table 4.4, the average score of all experts review evaluation was 3.66. The average scores of experts review in construct (instructional design), media (production), language, and content (subject matter) were 3.54, 3.84, 3.48, 3.79 respectively. Based on the validity categorization, the developed reading assessment was categorized as very highly valid.

4.1.3.3 One-to-one Evaluation

In the one-to-one evaluation, the students chosen from three different functional reading levels (independent, instructional, and frustration level) were administered to the product. They were given sixty minutes to complete 63 questions. After that, they were requested to fill out the questionnaire about the developed reading assessment. Those three students stated that they need extra time to answer all of the questions. Therefore, the writer revised the time allocation into 75 minutes. The information of the practicality questionnaire result of developed reading assessment was given below in which the average practicality score was 3.03 which categorized into high practicality (see table 4.5). The revised reading assessment after being evaluated in experts review and one-to-one evaluation was then called as prototype 2.

Table 4.5

Summary of Practicality in One-to-one Evaluation

No	Aspects	Average Score	Category
1	Instructional Clarity	3.00	High Practicality
2	Direction Clarity	2.89	High Practicality
3	Instructional Completeness	3.00	High Practicality

4	Instructional Effectiveness and Efficiency	3.27	Very High Practicality
5	Instructional Presentation	3.00	High Practicality
6	Typographical/Grammatical Error	3.00	High Practicality
	Total Average	**3.03**	**High Practicality**

4.1.3.4 Small Group Evaluation

The small group evaluation was intended to know the practicality of the product just like in one-to-one evaluation as well. The small group evaluation involved the nine students from three different functional reading levels (independent, instructional, and frustration). In this evaluation, the students stated that they need more extra time to finish the reading assessment. Then, the time allocation was revised into 90 minutes. Based on the result of the total 21 items of the questionnaires, the average score from the nine students was 3.15 which considered at high practicality level (see table 4.6). The revised reading assessment after being evaluated in small group evaluation was called prototype 3.

Table 4.6

Summary of Practicality in Small Group Evaluation

No	Aspects	Average Score	Category
1	Instructional Clarity	3.13	High Practicality
2	Direction Clarity	2.81	High Practicality
3	Instructional Completeness	3.15	High Practicality
4	Instructional Effectiveness and Efficiency	3.16	High Practicality
5	Instructional Presentation	3.30	Very High Practicality
6	Typographical/Grammatical Error	3.33	Very High Practicality
	Total Average	**3.15**	**High Practicality**

4.1.4.5 Field Test Evaluation

The field test evaluation was administered to class (44 students) that consist of the tenth, eleventh, and twelfth grade of SMK Kesehatan Athalla Putra Palembang. However, the students participating in one-to-one and small group evaluation were not involved anymore. The students were given the ESP reading assessment (prototype 3) consisting 70 multiple-choice questions which cover 8

aspects such as details, main idea, sequence, cause and effect, inference, critical reading, vocabulary, study skill – library.

The finding of the study showed that the product is potentially effective to be used for the nursing vocational students. It is showed from the students reading achievement after being tried out in the field test as 77% (categorized as good) of students passed the minimum score criterion which was 55. Based on the potential effect categorization, it can be said that the product had an average potential effect.

Furthermore, the students' score in the field test was used to measure the criterion-related validity of the developed ESP reading assessment. The students' score in ESP reading assessment was compared to the standardized test which is TOEIC. The statistics finding of the two set of scores had a positive strong correlation which is indicated by $r = 0.821$. The ESP reading assessment was also considered reliable as indicated by the value of Cronbach's Alpha obtained which was 0.882, meanwhile, the reliability of TOEIC was 0.795. Both of the tests were also considered normal, in which the developed reading assessment normality was 0.083 and the TOEIC normality was 0.110 in which significance value of the Shapiro – Wilk test is greater than 0.05.

4.2 Interpretation

The ESP reading assessment is developed for Nursing Vocational High School Students as to produce one of the alternative assessment devices in detecting the students' reading proficiency. This study adapted the developmental research proposed by Akker (1999) and the evaluation model proposed by Tessmer (1993) to ensure the validity, practicality, and potential effect of the developed product.

In developing ESP reading assessment, it is very crucial to consider the students' needs, students' functional reading level, and the readability level of the texts.

The nursing vocational high school students desired to have the specific topic (nursing) for learning input. Unfortunately, the reading materials and

assessments in their English textbooks are still general and do not cover their needs. In addition, the students stated that integrating the topics or contents related to the nursing field in reading texts is very important. The discrepancy between the students' needs and the reading materials/assessment in the students' textbooks based on instructional analysis implies that the reading assessments in vocational high schools do not really assess what should be assessed. It is assumed that the reading assessments for vocational high school students should cover the students' specific fields of interest.

Developing ESP reading assessment was needed in order to measure the students' achievement toward English in a specific field. In addition, it is what the English teachers should actually do and consider for the students in giving them the reading assessment. This condition is also in line with the study done by Diem, Mirizon, and Sitinjak (2018) which found health and medical care is one of the topics of interest the vocational students like to read.

Furthermore, although the high practicality average score resulted in one-to-one and small group evaluation, there is one aspect of practicality considered as the lowest aspect. It is direction clarity. The revision of direction clarity is especially done toward the information about the time allocation in one-to-one and small group evaluation. This implies that the time limitation is one of the considered aspects when giving the reading assessment to the students. This is in line with what Waugh and Gronlund (2013) stated that the directions for a test should be simple and concise yet contain information, such as (1) purpose of the test, (2) time allocation to complete the test, (3) how to record the answers, and (4) whether to guess when in doubt about an answer.

Then, the very high validity average score (3.66) resulted based on the four aspects (instructional design, media, language, and content) in experts review is assumed as a very good validity score of a product. However, since there are two experts (the second instructional design expert and content expert) stated that the developed reading assessment is valid with revision, it implies that the developed reading assessment was formally still needed to be improved to make it better, especially in test constructions and content in the prototype 1. This implies that to

produce a good quality of reading assessment, its development should be supported by significant others including instructional design expert and content specialist/expert in order to ensure a good validity result of the product. The different point of views among the experts also occured in the beginning, for example, the first instructional design expert stated the product as valid without revision meanwhile the second instructional design expert stated the product as valid with revision. However, both of the experts finally decided that the developed reading assessment was valid yet it should be revised based on experts' suggestions. Therefore, the product had been considered as a valid product to be evaluated for the next development phase.

Moreover, the developed product is potentially effective to be used for the target students. It can be seen from the students' reading achievement after being tried out in field test evaluation as 77 % (categorized as good mastery percentage) of the students passed the minimum score criterion which is 55. However, this result is not really satisfying enough because the students' reading achievement is still considered low due to the minimum score criterion which is just 55.

In line with the evaluation model proposed by Tessmer (1993), it can be said that the quality of the developed reading assessment has fulfilled the three criteria which are valid, practical and has the potential effect. However, some of the students still get a low achievement. Although the low achievement is achieved by some of the students, the developed reading assessment is assumed as good enough product since it can detect the students' strengths and weaknesses. This is in line with what Cooper, et al. (1988) stated that assessment is a process to find out about the students' strengths and weaknesses in reading which is as a tool used by the teacher to determine the students' needs, students' reading level and the best type materials for teaching reading. Furthermore, it can be assumed that although the experts stated that the developed reading assessment in this study is very highly valid, the developed reading assessment is considered medium-high practical according to the students.

In terms of criterion-related validity, the strong correlation (r = 0.821) resulted toward the correlation between the product and the standardized test

(TOEIC) means that the product can be used to measure the ESP reading achievement of nursing vocational high school students. Waugh and Gronlund (2013) stated that the criterion-related validity can be obtained by calculating the correlation coefficient between the two sets of measurements. Finally, it can be assumed that the developed reading assessment in this present study is good enough to be used as reading assessment for nursing vocational high school students.