

Chapter 4 Thesis

by Joko Slamet

Submission date: 15-Sep-2021 12:51PM (UTC+0700)

Submission ID: 1648882983

File name: 1988103009_Chapter_4-Joko_Slamet.docx (557.11K)

Word count: 6991

Character count: 35735

CHAPTER IV

FINDINGS AND DISCUSSION

This research resulted in the creation of a development product in the form of items analysis, which was named as @joe analysis. The product was created using Microsoft Excel as a base. The content of the product was fitted to the objectives of the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo's TOEFL-like placement test, and the product was developed in accordance with the results of the need analysis.

4.1 RESEARCH FINDINGS

4.1.1 Product Development

The need analysis was carried out by obtaining information from the administrator of the Master of English Education program at STKIP PGRI Sidoarjo as well as the Director of the Master of English Education program at STKIP PGRI Sidoarjo in relation to the TOEFL-like placement test that had been completed previously. The product development of the items analysis was organized based on the results of the needs analysis. The results of the need analysis revealed that the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo required the development of content related to the TOEFL-like placement test because the TOEFL-like has been conducted in the form of a paper-based placement test since 2017. Additionally, the test creators and team performed a manual scoring of the tests to ensure that

the results of the tests were accurate. As a result, the development of a product of the items analysis could be a solution to the problem in order to resolve it. The product of the development, called *@joe_analysis*, can be seen in the following figure.

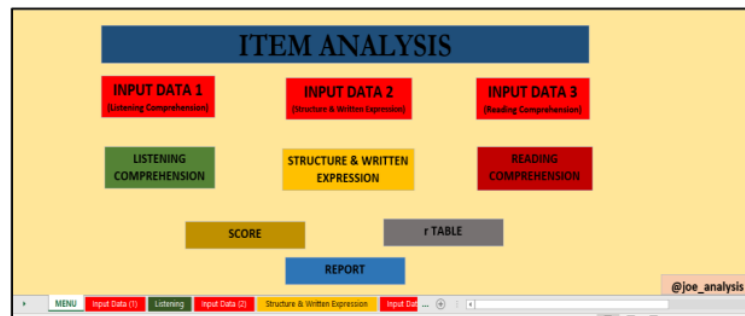


Figure 4.1 The Front Page of *@joe_analysis*

The product development of the items analysis was designed through Microsoft Excel base. In the front page, there are some buttons that can be used as the navigation to operate the items analysis. In first line, there are three buttons, namely: 'Item Analysis' as the title of the menu page, 'Input Data 1' which used for navigating the data of the test related to the 'Listening Comprehension' section, 'Input Data 2' as the navigation button for computing the data of the 'Structure & Written Expression' section, and 'Input Data 3' as the navigation button to input the data related to 'Reading Comprehension' section.

In the second line of *@joe_analysis*, there are three buttons provided namely: ¹ 'Listening Comprehension', 'Structure & Written Expression', and 'Reading Comprehension'. The button 'Listening

As seen in the figure 4.1, it can be described that the sheet of 'Input Data 1 is the are of product development in the listening comprehension section of the distribution of the answers. This sheet covers some data to be inputted to proceed the items analysis covering the identity of the test type, test held, number of questions, section name, the list of the test takers and the distribution of the answers in listening comprehension. The 'Menu' button on the right side above is used to navigate of to come to the main page in 'Menu' sheet.

The image shows a screenshot of an Excel spreadsheet titled "VALIDITY & ITEM ANALYSIS". The spreadsheet is designed for data entry, with a grid of 25 columns labeled "ITEMS" (numbered 1 to 25) and 25 rows labeled "STANDARDS" (numbered 1 to 25). The grid cells are currently empty, indicating it is a template for inputting data. At the bottom of the spreadsheet, there is a navigation bar with several buttons: "MENU", "Input Data (1)", "Listening", "Input Data (2)", "Structure & Written Expression", and "Input Data". A large, semi-transparent watermark reading "Page 1" is overlaid on the center of the spreadsheet grid.

Figure 4.3 The Sheet of Validity & Items Analysis of Listening Section

Figure 4.3 describes the sheet area of the 'Validity & Items Analysis' of the listening comprehension section. This sheet covers the data examine related to the results of the TOEFL-like listening section containing the validity of the items, reliability, index difficulty and index discrimination of the TOEFL-like items in listening comprehension section. The data are analyzed automatically based on the results inputted.

Figure 4.5 The Sheet of 'Structure & Written Expression'

The figure above showed the area of the data to be inputted in the 'Structure & Written Expression' section, specifies the sheet region of the 'Validity & Items Analysis' section of the 'Structure & Written Expression' section that contains the words "Validity & Items Analysis." 'Structure and Written Expression' is a TOEFL-like test that includes items that measure validity, reliability, index difficulty, and index discrimination. This sheet contains data on the validity of the items, reliability, index difficulty, and index discrimination of the TOEFL-like items in this section. The data is processed automatically in accordance with the results that have been entered.

		THE ANSWERS' DISTRIBUTION																																																					
		Distribution of Number of Items, Key Answers & Answers																																																					
No	Name of Students	NM	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50			
1	Student 1	1988103001	D	C	D	A	B	A	C	C	B	A	B	B	A	C	B	B	B	B	B	C	C	A	C	D	A	D	C	C	A	D	C	C	A	D	C	B	D	B	A	D	B	D	B	B	A	D	A						
2	Student 2	1988103002	D	C	D	D	B	D	A	C	B	C	A	B	A	D	C	B	A	D	B	A	C	B	A	C	D	B	A	A	C	B	A	A	C	B	A	A	B	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	
3	Student 3	1988103003	D	C	C	D	A	B	A	C	C	B	A	B	D	C	B	D	D	B	D	A	C	C	B	D	A	C	D	A	D	A	A	C	D	A	C	A	A	B	D	B	C	B	A	A	A	A	A	A	A	A	A	A	
4	Student 4	1988103004	A	D	D	D	D	A	A	B	A	B	A	C	A	D	C	A	A	C	D	C	A	A	C	A	D	D	A	D	A	A	C	D	D	A	A	A	A	A	D	A	D	A	C	D	A	A	A	A	A	A	A	A	A
5	Student 5	1988103005	D	C	B	D	D	A	C	C	B	C	B	A	D	C	D	B	A	D	B	D	A	B	D	B	C	D	B	A	A	C	B	D	B	A	A	C	B	D	A	A	B	D	C	D	C	B	A	C	A	D	B		
6	Student 6	1988103006	B	C	C	D	D	C	C	B	C	D	B	D	C	B	B	C	D	B	B	C	D	C	A	C	D	C	B	C	C	B	D	D	B	A	A	D	C	C	D	B	B	B	D	A	A	A	A	A	A	A	A		
7	Student 7	1988103007	D	D	D	D	D	C	A	C	B	A	B	C	D	B	B	C	D	B	C	D	B	C	A	D	C	D	D	D	D	C	D	C	D	A	B	C	B	C	B	C	B	B	C	B	C	B	C	B	C	B	C		
8	Student 8	1988103008	D	C	D	B	C	A	C	B	A	B	A	D	C	B	D	B	A	D	C	A	B	A	D	B	D	D	A	D	A	D	B	D	D	A	D	A	B	A	D	B	A	D	B	A	D	B	A	D	B	A	D		

Figure 4.6 The Sheet of 'Input Data 3'

It displayed the area of the distribution responses for the 'Reading Comprehension' component of the examination. The results of the test takers are displayed based on the information they provided. It is necessary to fill out this sheet with certain information in order to proceed with the items analysis. This includes information about the test type, when it was administered, how many questions were asked, what was covered in each section, who took the test, and the distribution of answers in the 'Reading Comprehension' section of the TOEFL-like test. The 'Menu' button on the right side of the screen above is used to navigate away from this main page in the 'Menu' sheet and back to the main page of @joe analysis, as shown in the figure above.

NO	STUDENTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	Student 1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	Student 2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	Student 3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	Student 4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	Student 5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	Student 6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	Student 7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	Student 8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	Student 9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	Student 10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	Student 11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	Student 12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	Student 13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	Student 14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	Student 15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	Student 16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	Student 17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	Student 18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	Student 19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	Student 20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 4.7 The Sheet of 'Reading Comprehension'

The amount of the data to be entered in the 'Reading Comprehension' section was depicted in the figure 4.7 indicates the sheet area of the 'Validity & Items Analysis' section of the 'Reading Comprehension' section that contains the words "Validity & Items Analysis." specifies the sheet location of the 'SReading Comprehension' section that contains the phrases "Validity & Items Analysis." A test of reading comprehension, similar to the TOEFL, has items that assess validity, reliability, difficulty indexing, and discriminating across indexes of difficulty. This sheet offers information about the validity of the items, their reliability, the index difficulty, and the index discrimination of the TOEFL-like items in this part, as well as other pertinent information. In accordance with the results that have been entered, the data is processed automatically by the system.

STKIP PGRI SDOARJO											
1											
2	Test : TOEFL										
4	Batch/Semester : 5/I										
5	Academic Year : 2019/2020										
6	Type of Test : Placement Test										
8	NO	STUDENTS	NIM	SECTION 1		SECTION 2		SECTION 3		TOTAL SCORE	RANK
				SCORE	CONVERTED	SCORE	CONVERTED	SCORE	CONVERTED		
10	1	Student 1	1988103001	26	48	28	52	39	54	613,33	6
11	2	Student 2	1988103002	39	57	34	58	40	55	666,67	2
12	3	Student 3	1988103003	37	55	26	50	33	50	616,67	5
13	4	Student 4	1988103004	37	55	24	48	39	54	623,33	4
14	5	Student 5	1988103005	30	51	25	49	36	52	606,67	7
15	6	Student 6	1988103006	35	54	28	52	36	52	626,67	3
16	7	Student 7	1988103007	33	52	22	46	34	51	606,67	9
17	8	Student 8	1988103008	32	52	24	48	33	50	600,00	8
18	9	Student 9	1988103009	45	62	36	61	43	58	693,33	1
19	10	Student 10	1988103010	28	49	21	45	32	49	476,67	10
20	11	Student 11	1988103011	28	49	20	44	33	50	476,67	10
21	12	Student 12	1988103012	25	48	21	45	30	48	470,00	14
22	13	Student 13	1988103013	27	49	18	42	31	48	463,33	15
23	14	Student 14	1988103014	26	48	18	42	31	48	460,00	16
24	15	Student 15	1988103015	30	51	19	43	32	49	476,67	10
25	16	Student 16	1988103016	25	48	21	45	32	49	473,33	13
26	17	Student 17	1988103017	23	47	19	43	26	45	450,00	18
27	18	Student 18	1988103018	23	47	20	44	26	45	453,33	17

Figure 4.8 The Converting of the TOEFL-like Score

The figure 4.8 above indicated that the results of the test takers' scores was displayed. This sheet contains some information related to the test' results of TOEFL-like. On the top area contains the identity of the institution of STKIP PGRI Sidoarjo, the information of the test held. The main area is the information toward the results of each section such as 'Section 1', 'Section 2', and 'Section 3' as well as the scores converted based on the TOEFL preparation taken from Longman by Debora Phillips as the sources of the tests' items. Additionally, this sheet also presents the ranks toward the grading of the test takers.

4.1.2 Items Analysis of the TOEFL-like Placement Test

The listening section, the structure and written expression section, and the reading section of the TOEFL-like placement test at STKIP PGRI Sidoarjo in the Master of English Education program (MPBI) are all comprised of three sections: listening, structure and written expression, and reading. The TOEFL Preparation of Longman,

written by Debora Philips, is used to double-check the classification of the assessed TOEFL-like test. Answering the second research question which was about the quality of the test items in the TOEFL-like placement test done by STKIP PGRI Sidoarjo, the classification was done into validity, reliability, index difficulty and discriminating power of the tests' items. Those are presented as follow:

4.1.2.1 Validity

As part of the product development, the validity of the formula correlation point biserial was calculated utilizing @joe analysis formula. It was determined that there were 18 students in the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo who served as the sample for field testing of the product. The accompanying table shows that there were 30 items classified as valid in the field testing of the product in listening section. As can be seen in the accompanying table, the results of the validity analysis of the item were satisfactory.

Table 4.1. The Validity of Listening Section

No	Validity Index	Item Question	Amount	Percentage
1	< 0.279 (invalid)	3, 5, 6, 9, 11, 12, 14, 18, 20, 24, 26, 28, 31, 32, 39, 40, 42, 45, 46, 47, 50	20	40%
2	≤ 0.279 (valid)	1, 2, 4, 7, 8, 10, 13, 15, 16, 17, 19, 21, 22, 23, 25, 27, 29, 30, 33, 34, 35, 36, 37, 38, 41, 43, 44, 48, 49	30	60%

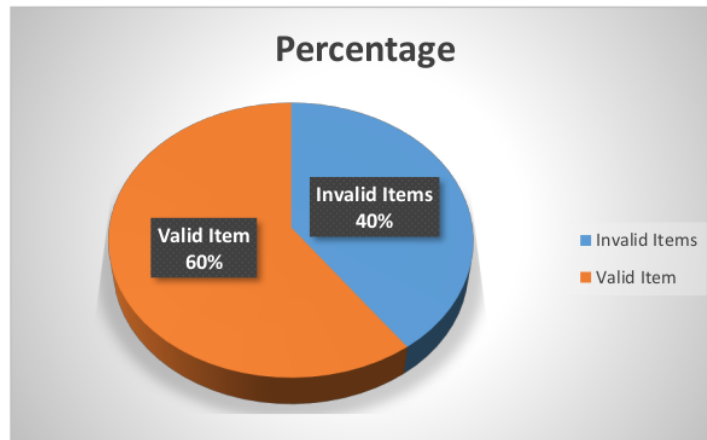


Figure 4.9 Percentage of Validity Item of Listening Section

As part of the product development process, the validity of the formula correlation point biserial was determined by applying the @joe analysis formula and a correlation point biserial. In order to conduct field testing of the product, it was established that there were 18 students enrolled in the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo who acted as a sample for the product. In the 'Structure & Written Expression' section of the accompanying table, it is shown that there were 30 elements classified as valid in the field testing of the product. According to the information provided in the accompanying table, the item's validity analysis yielded acceptable results.

Table 4.2. The Validity of Structure & Written Expression Section

No	Validity Index	Item Question	Amount	Percentage
----	----------------	---------------	--------	------------

1	< 0.279 (invalid)	1, 3, 4, 6, 7, 9, 10, 13, 14, 16, 18, 19, 20, 21, 24, 25, 26, 32, 33, 37, 38, 40	22	55%
2	≤ 0.279 (valid)	2, 5, 8, 11, 12, 15, 17, 22, 23, 27, 28, 29, 30, 31, 34, 35, 36, 39	18	45%

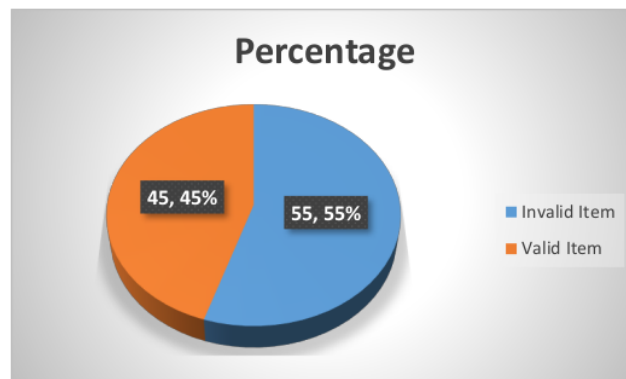


Figure 4.10 Percentage of Validity Item of Structure & Written Expression Section

The validity of the correlation point biserial formula was determined as part of the product development process by applying the @joe analysis formula and a correlation point biserial as part of the product development process. For field testing of the product, it was determined that there were 18 students enrolled ³ in the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo who served as a sample. When looking at the 'Reading Comprehension' portion of the accompanying table, it can be seen that there were ¹ 30 items that were classified as valid during the product's field testing. According on the information presented in the accompanying

table, the item's validity analysis produced results that were acceptable to the test takers.

Table 4.3. The Validity of Reading Comprehension Section

No	Validity Index	Item Question	Amount	Percentage
1	< 0.279 (invalid)	2, 6, 7, 9, 11, 12, 13, 16, 18, 20, 21, 24, 28, 29, 30, 31, 32, 35, 36, 37, 38, 41, 42, 44, 45, 46, 47, 50	27	54%
2	≤ 0.279 (valid)	1, 3, 4, 5, 8, 10, 14, 15, 17, 19, 22, 23, 25, 26, 27, 33, 34, 39, 40, 43, 48, 49	23	46%

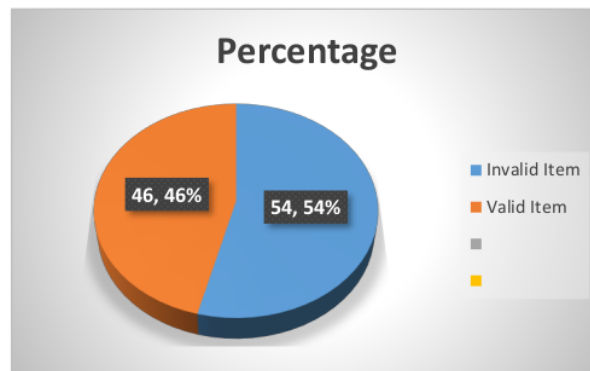


Figure 4.11 Percentage of Validity Item of Reading Comprehension Section

4.1.2.2 Reliability

The results of the reliability of the items test using @joe_analysis was based on the standard that if r11 is less than 0.279, the things that were previously considered to be reliable are now considered to be less reliable. If r11 is less than 0.279,

it indicates that a question used to be unreliable or had a low reliability level. The items tested for the 'Listening Section' had a reliability index of 0.887, which was used in the calculation of the overall reliability. This implies that the items test of listening comprehension component of the TOEFL-like placement test at STKIP PGRI Sidoarjo was assessed as having "Very Good Reliability".

Using @joe analysis, the results of the reliability of the items test were compared to a standard that states that if r_{11} is less than 0.312, things that were previously thought to be trustworthy are now thought to be less dependable. If the reliability coefficient (r_{11}) is less than 0.312, it implies that a question was previously unreliable or had a low level of reliability. In the 'Structure & Written Expression' section, the items that were examined had a reliability index of 0.999, which was used to determine the overall dependability of the section. Thus, the 'Structure and Written Expression' component of the TOEFL-like placement test at STKIP PGRI Sidoarjo was found to have "Excellent Reliability" in terms of the items tested.

After doing a @joe analysis, the findings of the reliability of the items test were contrasted with a criterion that specifies, when the correlation coefficient is less than 0.279, things that were previously deemed to be trustworthy are now considered to

be less trustworthy. It is possible that a question was previously unreliable or had a low level of reliability if its reliability coefficient (r_{11}) is less than 0.279. During the examination of the 'Reading Comprehension' section, the items under consideration received a reliability index of 0.753, which was used to assess the section's overall trustworthiness as a whole. This means that when it comes to the things assessed, the "Reading Comprehension" component of the TOEFL-like placement examination at STKIP PGRI Sidoarjo was judged to have "Good Reliability," according to the results.

4.1.2.3 Index Difficulty

For the purposes of interpretation, classification was used to group questions into categories. The results of the calculation of the level of difficulty were divided as follows: 0,00-0,29 for category difficulty questions; 0,30-0,69 for category medium questions; 0,70-1,00 for category easy questions. After conducting an investigation through @joe analysis it was discovered that the tough category contained 28 items (56%), the medium category contained 22 items (44%), and the easy category contained just 0 items (0%) based on the results of the investigation. The following is a list of the difficulty levels in items analysis of 'Listening Comprehension' section:

Table 4.4. The Distribution of Difficulty Index of Listening Comprehension Section

No	Difficulty Index	Item Question	Amount	Percentage
1	0,00 – 0,29 (difficulty)	2, 3, 5, 8, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 27, 30, 34, 38, 41, 42, 43, 45, 46, 50	22	44%
2	0,30 – 0,69 (medium)	1, 4, 6, 7, 9, 12, 24, 26, 28, 29, 31, 32, 33, 35, 36, 37, 39, 40, 44, 47, 48, 49	28	56%
3	0,70 – 1,00 (easy)	-	0	0%

Items' question categories were created for the purpose of interpretation by classifying them according to their content. The following are the divisions of the ² results of the computation of the level of difficulty: a. Category difficulties questions range from 0,00 to 0,29; category medium questions range from 0,30-0,69; and category easy questions range from 0,70 to 1,00. According to the findings of the inquiry, the challenging category contained 28 items (70%) in the 'Structure & Written Expression' area, the medium category contained 22 items (30%), and the easy category contained zero items (0%). The following is a breakdown of the different levels of difficulty in item analysis:

Table 4.5. The Distribution of Difficulty Index of Structure & Written Section

No	Difficulty Index	Item Question	Amount	Percentage
1	0,00 – 0,29 (difficulty)	4, 5, 8, 10, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 27, 28, 29,	28	70%

		30, 31, 32, 34, 35, 36, 37, 38, 40		
2	0,30 – 0,69 (medium)	1, 2, 3, 6, 7, 9, 11, 19, 25, 26, 33, 39	12	30%
3	0,70 – 1,00 (easy)	-	0	0%

It was decided to construct test's items categories for the purpose of interpretation by categorizing questions according to their substance of difficulty index. The following are the divisions of the ² results of the computation of the level of difficulty, as determined by the results of the computation: A. Category tough questions vary from 0,00 to 0,29; category medium questions range from 0,30 to 0,69; and category easy questions range from 0,70 to 1,00 in difficulty. Following the findings of the investigation, the tough category had 21 items (42%) in the 'Reading Comprehension' section, the medium category contained 22 items (58%), and the easy category contained zero items (0%). Item analysis was divided into three categories of difficulty, as follows: easy, medium, and difficult.

4.1.2.4 Discriminating Power

A categorization was ² used to interpret the results of the calculation of the 'Discrimination Index,' with values ranging from 0,00 to 0,19 falling into the category of 'Poor'; 0,20 to 0,39 falling into the category of 'Enough'; 0,40-0,69 falling into the

category of 'Good,' and 0,70-1,00 falling into the category of 'Very Good.' Following the results of the analysis undertaken by @joe analysis, it was discovered that items of the TOEFL-like test in the 'Listening Comprehension' section scored in the 'Poor' category of the 'Discrimination Index' totaled 27 items (54%), 16 items (32%), 7 items (14%), and 0 item (0%) went on to score in the 'Very Good' category of the 'Discrimination Index'. On the basis of the discrimination index, the following 50 items were presented:

Table 4.6. The Distribution of Discrimination Index of Listening Comprehension Section

No	Discrimination Index	Item Question	Amount	Percentage
1	0,00-0,19 (Poor)	1, 2, 3, 6, 7, 8, 9, 10, 13, 16, 20, 21, 23, 24, 25, 27, 29, 30, 31, 36, 38, 41, 42, 43, 45, 46, 50	27	54%
2	0,20-0,39 (Enough)	4, 5, 15, 16, 20, 22, 28, 32, 33, 34, 39, 40, 44, 47, 48, 49	16	32%
3	0,40-0,69 (Good)	14, 17, 18, 26, 35, 37	7	14%
4	0,70-1,00 (Very Good)	-	0	0%

The results of the calculation of the 'Discrimination Index' were categorized, with values ranging from 0,00 to 0,19 falling into the category of 'Poor,' 0,20 to 0,39 falling into the category of 'Enough,' 0,40-0,69 falling into the category of 'Good,' and values ranging from 0,70-1,00 falling into the category of 'Very

Good,' respectively. Following the results of the analysis undertaken by @joe analysis, it was discovered that items of the TOEFL-like test in the 'Structure & Written Expression' section scored in the 'Poor' category of the 'Discrimination Index' totaled 27 items (54%), 16 items (32%) in the 'Enough' category, 7 items (14%) of 'Good' category, and 0 item (0%) went on to score in the 'Very Good' category of the 'Discrimination Index'. The following 40 items of the 'Structure & Written Expression' comprehension portion were provided on the basis of the discriminating index; they are listed below:

Table 4.7. The Distribution of Discrimination Index of Structure & Written Expression Section

No	Discrimination Index	Item Question	Amount	Percentage
1	0,00-0,19 (Poor)	1, 4, 5, 6, 8, 9, 10, 13, 16, 19, 21, 23, 24, 25, 27, 29, 30, 31, 32, 36, 38	21	55%
2	0,20-0,39 (Enough)	2, 3, 7, 11, 14, 15, 20, 22, 28, 33, 34, 37, 39, 40	14	35%
3	0,40-0,69 (Good)	12, 17, 18, 26, 35	5	10%
4	0,70-1,00 (Very Good)	-	0	0%

Using the results of the calculation of the 'Discrimination Index,' the results were divided into four categories: poor, enough, good, and very good. Values ranging from 0,00 to 0,19 falling into the category of 'Poor'; 0,20 to 0,39 falling into the

category of 'Enough'; 0,40-0,69 falling into the category of 'Good,' and 0,70-1,00 falling into the category of 'Very Good.'. Following the results of the analysis undertaken by @joe_analysis, it was discovered that items of the TOEFL-like test in the 'Reading Comprehension' section scored in the 'Poor' category of the 'Discrimination Index' totaled 34 items (68%), 13 items (26%) in the 'Enough' category, 3 items (6%) of 'Good' category, and 0 item (0%) went on to score in the 'Very Good' category of the 'Discrimination Index'. In response to the discriminating index results, the following 50 items from the 'Reading Comprehension' section; they are listed below.

Table 4.8. The Distribution of Discrimination Index of Reading Comprehension Section

No	Discrimination Index	Item Question	Amount	Percentage
1	0,00-0,19 (Poor)	1, 3, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 20, 21, 23, 24, 25, 27, 29, 30, 31, 32, 36, 38, 40, 41, 42, 44, 47, 48, 49, 50	34	68%
2	0,20-0,39 (Enough)	2, 4, 11, 18, 19, 22, 26, 28, 33, 34, 37, 39, 43	13	26%
3	0,40-0,69 (Good)	12, 17, 35	3	6%
4	0,70-1,00 (Very Good)	-	0	0%

4.2 DISCUSSION

4.2.1 Product Development

In conducting test, STKIP PGRI Sidoarjo's TOEFL-like test, which is identical to the TOEFL in terms of scoring and question creation, is similar to that test in that it uses the TOEFL as a benchmark for comparison. The test items for the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo are not created by the program; rather, the test creators use questions from other sources, such as the TOEFL test from Longman by Debora Philips. It is separated into ¹¹ three sections: listening comprehension, structure and written expression, and reading comprehension. Every new student is required to take this test as part of their requirement to graduate from the MPBI program and become a candidate for a master's degree. The passing score on this test is 475 points. After failing the first test, the test takers can attempt the second and subsequent tests until they achieve the desired score on all of the tests taken. One of the requirements for participation in the thesis examination and graduating with a master's degree is a certificate from a test that is similar to the TOEFL.

According to the results of a TOEFL-like placement test that had been completed previously, information from the administrator of the Master of English Education program at STKIP PGRI Sidoarjo as well as the Director of the Master of English Education program at STKIP PGRI Sidoarjo was obtained for the purpose of conducting a need analysis. In order to respond to the findings of the needs analysis, a product development plan for the items analysis was put together. The

results of the need analysis revealed that the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo required the development of content related to the TOEFL-like placement test because the TOEFL-like placement test has been conducted in the form of a paper-based placement test since 2017. The content related to the TOEFL-like placement test was developed and implemented in the following semester. Additional manual scoring of the tests was carried out by the test creators and team in order to confirm that the findings of the tests were correct. Because of this, developing a product from the items analysis could be a solution to the problem that will allow it to be resolved in the big scheme of things. The creation of a product known as @joe analysis is now underway.

The product development of the items analysis was planned using a Microsoft Excel database as a starting point. To operate the items analysis, there are certain buttons on the front page that may be used as navigation to operate the items analysis. There are three buttons in the first line, which are as follows: 'Item Analysis' as the title of the menu page, 'Input Data 1' which is used to navigate the data of the test related to the 'Listening Comprehension' section, 'Input Data 2' which is used to compute the data of the 'Structure & Written Expression' section, and 'Input Data 3' which is used to input the data related to the 'Reading Comprehension' section.

There are three buttons offered in the second line of @joe analysis, which are as follows: 'Listening Comprehension', 'Structure & Written Expression', and 'Reading Comprehension'. It is necessary to use the button 'Listening Comprehension' in order to travel between outcomes of tests in the listening comprehension section of the TOEFL-like test. When the part is finished, the button 'Structure & Written Expression' is used to browse around it in terms of structure and written expression. Finally, the option titled "Reading Comprehension" is utilized to scroll through the results of the TOEFL-like reading comprehension component of the test.

"Score" and "r table" are the names of two buttons that appear in the third line. They perform entirely distinct functions when these two buttons are pressed in succession. A purpose of the "Score" button is to navigate through the sheet, which displays the results of the overall score and the converted score in the TOEFL-like format. The "r Table" button, on the other hand, is utilized to display a list of the correlation product moments that were used in computing the data in the @joe analysis during this time.

The page of 'Input Data 1' also contains information on product development, which is found in the listening comprehension section, as well as information about distribution of the answers. There are some pieces of information that need to be entered in order to proceed with the items analysis. These include information about the test type, the

date of the test, the number of questions asked, the name of the section, the list of those who took the test, and the distribution of the answers in listening comprehension. The 'Menu' button on the right side of the screen above is used to go out of the 'Menu' sheet and back to the main page.

Afterwards, there is a sheet area for the 'Validity & Items Analysis' of the part on listening comprehension. Contains information ⁴ on the validity of the items, the reliability of the items, index difficulty and index discrimination of the TOEFL-like items in the listening comprehension part, as well as the reliability of the TOEFL-like items in the listening comprehension section. The data is examined automatically in accordance with the results that have been entered.

the area of the distribution replies in the section titled "Structure & Written Expression" The results of the test takers are displayed based on the information they provided. It is necessary to fill out this sheet with certain information in order to proceed with the items analysis. This includes information about the test type, when it was administered, how many questions were asked, what was covered in each section, who took the test, and the distribution of answers ¹ in the 'Structure & Written Expression' section of the TOEFL-like test. This sheet also contains information about the test takers. The 'Menu' button on the right side of the screen above is used to navigate away from this main page in the

'Menu' sheet and back to the main page of @joe analysis, as shown in the screenshot above.

It also specifies a sheet region within the 'Structure & Written Expression' section that contains the words 'Validity & Items Analysis,' as well as the area of data to be inputted in that section. Additionally, it specifies the sheet region within the 'Structure & Written Expression' section that contains the words 'Validity & Items Analysis,' as well as the area of data to be inputted in that section. It is similar to the TOEFL in that it includes items that measure validity, reliability, index difficulty, and index discrimination. It is administered online. This sheet offers information about the validity of the items, their reliability, the index difficulty, and the index discrimination of the TOEFL-like items in this part, as well as other pertinent information. The information is processed automatically in accordance with the search results that have been submitted into the system.

Afterwards, the region of the distribution replies for the 'Reading Comprehension' component of the examination was presented on the computer screen. Participants' test results are displayed based on the information they gave throughout the assessment. In order to proceed with the item analysis, it is important to complete this sheet with specific information. Included in this report is information regarding the test type, when it was delivered, how many questions were asked, what was covered in each section, how many people participated, and the

distribution of answers in the 'Reading Comprehension' section of the TOEFL-style exam. By clicking on the 'Menu' button on the right-hand side of the screen above, you can travel away from this main page and back to the main page of @joe analysis.

The amount of data to be entered in the 'Reading Comprehension' section was then depicted, with the sheet area of the 'Validity & Items Analysis' section of the 'Reading Comprehension' section that contains the words "Validity & Items Analysis" being indicated. specifies the sheet location of the 'SReading Comprehension' section that contains the phrases "Validity & Items Analysis." follows. A reading comprehension test, comparable to the TOEFL, contains items that assess validity, reliability, difficulty indexing, and differentiating between different indexes of difficulty, among other things. On this sheet, you will find information regarding the items' validity as well as their reliability, as well as their index difficulty and index discrimination for the TOEFL-like items in this section, among other useful details. The data is processed automatically by the system in accordance with the results that have been input.

After that, the results of the test takers' scores were revealed on screen. This document offers some information on the outcomes of the TOEFL-like test and how they were obtained. The top section contains the name of the institution, STKIP PGRI Sidoarjo, as well as the date and time of the test. Section 1 contains information about the results of

each section, such as 'Section 2' and 'Section 3', as well as the scores converted based on the TOEFL preparation taken from Longman by Debora Phillips as the sources of the tests' items. Section 2 contains information about the results of each section, such as 'Section 1' and 'Section 2'. Additionally, this sheet includes the test takers' rankings in relation to the overall grading of the test.

4.2.2 Items Analysis of the TOEFL-like Placement Test

¹ It is divided into three sections: listening, structure and written expression, and reading. The listening section, ¹² structure and written expression section, and reading section of the TOEFL-like placement test at STKIP PGRI Sidoarjo in the Master of English Education program (MPBI) are all made up of ⁶ three sections: listening, structure and written expression, and reading. It is necessary to double-check the classification of the assessed TOEFL-like test using the Longman TOEFL Preparation of Debora Philips, which was written by Debora Philips. In order to answer the second study question, which was regarding the quality of the test items in the TOEFL-like placement test developed by STKIP PGRI Sidoarjo, the items were classified into four categories: validity, reliability, index difficulty, and discriminating power.

As part of the product development process, the validity of the formula correlation point biserial was determined by applying the @joe analysis formula and a correlation point biserial. In order to conduct

field testing of the product, it was established that there were 18 students enrolled in the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo who acted as a sample for the product. There were 30 things classified as valid in the listening section of the product field testing, as shown in the table that is included with this document: The validity examination of the item yielded satisfactory results, as may be seen in the accompanying table.

The validity of the correlation point biserial formula was determined as part of the product development process by applying the @joe analysis formula and a correlation point biserial as part of the product development process. For field testing of the product, it was determined that there were 18 students enrolled in the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo who served as a sample. When looking at the 'Structure & Written Expression' portion of the accompanying table, it can be seen that there were 30 elements that were identified as legitimate during the product's field testing. According on the information presented in the accompanying table, the item's validity analysis produced results that were acceptable to the test takers

Using the @joe analysis formula and a correlation point biserial as part of the product development process, it was possible to determine the validity of the correlation point biserial formula. This was done as part of the product development process. According to preliminary

research, there were 18 students enrolled in the Master of English Education program (MPBI) at STKIP PGRI Sidoarjo who were used as a sample for field testing the product. In the accompanying table, it can be seen that there were 30 items that were classified as legitimate during the product's field testing, which can be found in the section titled "Reading Comprehension." In accordance with the information shown in the accompanying table, the item's validity analysis yielded results that were deemed acceptable by the test takers during the administration.

Using @joe analysis, the results of the reliability of the items test were compared to a standard that states that if r_{11} is less than 0.279, things that were previously regarded as dependable are now regarded as less reliable. If the correlation coefficient (r_{11}) is less than 0.279, it suggests that a question was previously unreliable or had a low level of dependability. The dependability index for the items tested for the 'Listening Section' was 0.887, which was used in the calculation of the overall reliability. This suggests that the items test of hearing comprehension component of the TOEFL-like placement test at STKIP PGRI Sidoarjo was found to be of "Very Good Reliability," which means that it was found to be highly reliable.

The results of the reliability of the items test were compared to a criterion that indicates that if r_{11} is less than 0.312, things that were previously considered trustworthy are now considered less dependable.

It is possible that a question was previously unreliable or had a poor level of consistency if the reliability coefficient (r_{11}) is less than 0.312. The items that were analyzed in the 'Structure and Written Expression' section had a reliability index of 0.999, which was used to determine the overall dependability of the section. It was determined that in terms of the items evaluated, the "Structure and Written Expression" component of the TOEFL-like placement examination at STKIP PGRI Sidoarjo had "Excellent Reliability," which means it was highly reliable.

A @joe study was carried out to compare and contrast those findings with a criterion that states that when the correlation coefficient is less than 0.279, things that were previously thought to be trustworthy are now considered to be less trustworthy. If the reliability coefficient (r_{11}) of a question is less than 0.279, it is conceivable that the question was previously faulty or had a poor degree of reliability before being tested. After a thorough assessment of the 'Reading Comprehension' section, it was determined that the items under consideration had an overall reliability index of 0.753, which was used to determine the overall trustworthiness of the section as a whole. In terms of the things that were tested, the "Reading Comprehension" component of the TOEFL-like placement examination at STKIP PGRI Sidoarjo was found to have "Good Reliability," according to the results of the study.

Question categories were created for the purpose of interpretation by classifying them according to their content. The following are the divisions of the ² results of the computation of the level of difficulty: a. Category difficulties questions range from 0,00 to 0,29; category medium questions range from 0,30-0,69; and category easy questions range from 0,70 to 1,00. According to the findings of the inquiry, the challenging category comprised 28 items (56 percent), the medium category contained 22 items (44 percent), and the easy category contained zero items (0 percent).

It was decided to categorize items into question categories for the purpose of interpretation by classifying them according to their content. As a result of the computation of difficulty, questions in the following categories are divided into three groups: a. Category difficulties questions are scored between 0,00 and 0,29; category medium questions are scored between 0,30-0,69; and category easy questions are scored between 0,70 and 1,00. A study conducted by a private company found that the tough category comprised 28 items (70 percent) in the 'Structure & Written Expression' section, the medium category featured 22 things (30 percent), and the simple category contained no items (0 percent).

According to the substance of difficulty index, it was determined to establish test item categories for the purpose of interpretation by categorizing questions according to their difficulty index. In accordance with the computation findings, the following are the levels of difficulty,

which are divided into three divisions: easy, medium, and severe. A. The difficulty of category tough questions ranges from 0,00 to 0,29; the difficulty of category medium questions is from 0,30 to 0,69; and the difficulty of category light questions ranges from 0,70 to 1,00. After conducting an examination, it was discovered that the difficult category had 21 things (42 percent) in the 'Reading Comprehension' part, the medium category had 22 items (58 percent), and the easy category had zero items in the section (0 percent). In order to differentiate between easy, medium, and difficult, the item analysis was separated into three categories of difficulty: easy, medium, and difficult.

The results of the calculation of the 'Discrimination Index' were categorized, with values ranging from 0,00 to 0,19 falling into the category of 'Poor,' 0,20 to 0,39 falling into the category of 'Enough,' 0,40-0,69 falling into the category of 'Good,' and values ranging from 0,70-1,00 falling into the category of 'Very Good,' respectively. The results of the analysis conducted by @joe analysis revealed that 27 items of the TOEFL-like test in the 'Listening Comprehension' section scored in the 'Poor' category of the 'Discrimination Index,' 16 items scored in the 'Poor' category of the 'Discrimination Index,' 7 items scored in the 'Poor' category of the 'Discrimination Index,' and 0 item (0 percent) scored in the

Using the results of the calculation of the 'Discrimination Index,' the results were divided into four categories: poor, enough, good, and very

good. with values ranging from 0,00 to 0,19 falling into the category of 'Poor,' 0,20 to 0,39 falling into the category of 'Enough,' 0,40-0,69 falling into the category of 'Good,' and values ranging from 0,70-1,00 falling into the category of 'Very Good,' respectively. According to the findings of the @joe analysis's analysis, items ¹³ of the TOEFL-like test in the 'Structure & Written Expression' section scored in the 'Poor' category of the 'Discrimination Index' totaled 27 items (54 percent), 16 items (32 percent) in the 'Enough' category, 7 items (14 percent) in the 'Good' category, and 0 item (0 percent) went on to score in the 'Vital'

The findings of the computation of the 'Discrimination Index' were grouped into four groups based on the outcomes of the calculation: poor, sufficient, good, and very good. From 0,00 to 0,19, values are classified as 'Poor,' 0,20 to 0,39, as 'Enough,' 0,40-0,69 are classified as 'Good', and from 0,70-1,00 are classified as "Very Good." According to the findings of the @joe analysis's analysis, items of the TOEFL-like test in the 'Reading Comprehension' section scored in the 'Poor' category of the 'Discrimination Index' totaled 34 items (68 percent), 13 items (26 percent) in the 'Enough' category, 3 items (6 percent) in the 'Good' category, and 0 item (0 percent) went on to score in the 'Very Good' category.

It is critical that the teaching materials generated by the researcher be vetted by specialists before they can be implemented in the classroom by the researcher. Following the experts' statement, the

review was carried out to ensure that the teaching materials were well created and ready to be used, and that the experts' statement was followed by the experts' statement. The results of the evaluation were taken from the observation sheets that had been provided to the experts, and they were quoted in full. In addition to the systematic content of the materials, formulated statistics based ⁴ on the validity, reliability, index difficulty and discriminating index theories utilized, and drawings that have already been validated with a 'A Good Mark' were included on the observation sheet. The product development team at @joe_analysis claimed 'Valid' once all of the items were marked and the results were positive. This was in accordance with the experts' suggestions: This is due to the fact that the materials have been meticulously formulated for the reasons listed below: The aims and goals have been outlined, a development framework has been established, and the concept of integrated skills is being presented in a methodical manner on the basis of a TOEFL-like test. After all of the things were marked in the same way by the other experts and the results were excellent, the product development materials were found to be valid in accordance with the experts' recommendations: It is the desire of the English language specialist to implement the TOEFL-like test at STKIP PGRI Sidoarjo, and it might also serve as inspiration for the development of a proper TOEFL-like test in the future.

Chapter 4 Thesis

ORIGINALITY REPORT

6%

SIMILARITY INDEX

5%

INTERNET SOURCES

3%

PUBLICATIONS

1%

STUDENT PAPERS

PRIMARY SOURCES

1

ejournal.iainbukittinggi.ac.id

Internet Source

2%

2

studylib.net

Internet Source

1%

3

Madhu Neupane. "Should we Bend towards Blending? How?", Journal of NELTA, 2018

Publication

1%

4

digilib.uinsby.ac.id

Internet Source

<1%

5

Abdulloh Abdulloh, Sarsono Sarsono, Slamet Basuki. "PREPARING PBT TOEFL PREDICTION: AN EXPERIENCE OF ONLINE TEACHING TOEFL PREPARATION", Journal of English Language and Literature (JELL), 2021

Publication

<1%

6

www.jurnal.unsyiah.ac.id

Internet Source

<1%

7

Hale, Gordon A., Charles W. Stansfield, and Richard P. Duran. "SUMMARIES OF STUDIES INVOLVING THE TEST OF ENGLISH AS A

<1%

FOREIGN LANGUAGE, 1963-1982", ETS
Research Report Series, 1984.

Publication

8	repository.phb.ac.id Internet Source	<1 %
9	www.manhattanreview.com Internet Source	<1 %
10	Submitted to Macquarie University Student Paper	<1 %
11	en.apu.ac.jp Internet Source	<1 %
12	herukeren-herusetiawan.blogspot.com Internet Source	<1 %
13	Syaefani Arif Romadhon, Yuvita Yuvita. "Toefl Result of Politeknik Harapan Bersama Tegal Student's Majoring Mechanical Engineering", English Focus: Journal of English Language Education, 2018 Publication	<1 %
14	Reed, D.J.. "The relationship between criterion-based levels of oral proficiency and norm-referenced scores of general proficiency in English as a second language", System, 199208 Publication	<1 %

Exclude quotes On

Exclude matches Off

Exclude bibliography On