POLYGLOT: Jurnal Ilmah DOI: http//dx.doi.org/10.1966/pji.v18i1.4908
Vol 18, No.1 (2022): Januari 2022 page : 1-18 P-ISSN : 1907-6134. E-ISSN: 2549-1466

COMPREHENSION PROCESSES AND CRITICAL THINKING SKILLS IN TWO UNIVERSITY ENTRANCE EXAMS IN INDONESIA: A CONTENT ANALYSIS OF SELEKSI BERSAMA MASUK PERGURUAN TINGGI NEGERI AND SELEKSI MASUK UI 2008-2019

Bernard Richard Nainggolan[1], Pupung Purnawarman[2], Didi Sukyadi[3]
[1) 2) 3)]Universitas Pendidikan Indonesia
bernardnainggolan@upi.edu[1], purnawarman@upi.edu[2], dsukyadi@upi.edu[3]

ABSTRACT

The entrance exams, particularly the English questions, are believed to qualify candidates' comprehension and critical thinking skills. Therefore, the investigation to university exams in Indonesia became significantly important. As a matter of fact, studies on entrance exams were rare and the trends of English questions in the exams were known a little. Thus, this study aimed at investigating the trends of comprehension processes and higher-order thinking skills manifested in both the Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) and the Seleksi Masuk UI (SIMAK UI) within a decade (2009-2018). Employing a mixed-method of content analysis, the questions were investigated quantitatively and qualitatively to result in several findings, such as 1) types of questions tested, 2) comprehension processes, 3) and higher-order thinking questions, and 4) construction MCQs to promote critical thinking skills. Finally, implications of the findings were then elaborated.

Keywords: university entrance exams, higher-order thinking, comprehension processes, MCQs

Introduction

To qualify candidates, university entrance exams become the tools to test their student candidates with expected skills and competencies. One of the expected skills and competencies in universities are the critical thinking skills. Therefore, critical thinking skills have, undoubtedly, become one of the most fundamentally demanded skills to support students' success in academic ecosystem of the 21-st century (NACE, 2019; P21CS, 2008; PPRC, 2010)**.** In relation to critical thinking skills, university entrance exams are conducted to select the best candidates with good academic competence and, of course critical, thinking skills. In the university entrance exams, the language tests are conducted with the assumption that the score tests will deliver real-life consequences, such as access to a valued position, service, or status (Deygers et al., 2018). This means good entrance exams need to consider the social context and consequences.

The urgency of English tests in entrance exams is to facilitate the first year university students to skills needed such as expressing ideas accurately, understanding coherence and cohesion, taking class notes, composing a logical argumentation, having grammatical accuracy, and summarizing long text (Deygers et al., 2018). As the matter of fact, the English tests of university entrance in Indonesia are still concerned with the comprehension. Despite the high demand of critical thinking in higher education, especially in universities, it is assumed that the critical questions are also to be found in the entrance exams.

There are less studies on university entrance exams, particularly English entrance exams. In Malaysia, the Predictive Validity of the Malaysian University English Test (MUET), as a measure of students' proficiency, prerequisite for admission, and placement in various academic program, was examined (Rethinasamy & Chuah, 2011; Samad, 2008). The study indicated that there was a significant relationship between undergraduate MUET scores and their grades in a language course. Another study related to entrance exam is a study of Paribakht and Webb (2016) that investigated the academic vocabulary presence in English proficiency test for admission purposes at Canadian universities. This studied examined the presence of vocabulary of Academic World List (AWL) and the vocabulary was consistently present and substantial in

academic passage with below established levels. Hemati and Baghaei (2020) studied the English reading comprehension section of the Iranian University Entrance Exam (IUEE). Using a Cognitive Diagnostic Models (CDMs), they found information related to test takers' weaknesses in the subskills and subprocesses of reading. Using the GDINA model, one of CDMs models, the study found that skills that test takers were troubled with are 'making inferences', evaluating response options', and 'extracting explicit information'. The studies aforementioned have not attempted to investigate how critical thinking skills are manifested in the items of tests and comprehension processes are represented in the English entrance exams. Therefore, this study is to fill in the gap in English entrance exams, that is the investigation on the critical thinking and comprehension processes in the entrance exams.

The formulated research questions of this research are:

a. How do English questions of SBMPTN and SIMAK UI differ in terms of types of questions?
b. How do reading tests in both SBMPTN and SIMAK UI differ in terms of comprehension processes?
c. How do English questions of SBMPTN and SIMAK UI differ in terms of learning objectives of Revised Bloom's taxonomy?
d. How do MCQs in SBMPTN and SIMAK UI promote higher-order thinking?

Theoretical Framework

Revised Bloom's taxonomy

The study applied Revised Bloom's Taxonomy with a two-dimensional domain, cognitive and knowledge dimensions (Anderson & Krathwohl, 2001; Bloom, 1956; Krathwohl, 2008).

Table 1. Bloom's revised structure of the knowledge dimension (Krathwohl, 2008).

A. Factual knowledge	Aa. Knowledge of terminology
	Ab. Knowledge of specific details and elements
B. Conceptual Knowledge	Ba. Knowledge of classification and categories
	Bb. Knowledge of principles and generalizations
	Bc. Knowledge of theories, models, and structures
	Ca. Knowledge of subject-specific skills and algorithms

C. Procedural Knowledge	Cb. Knowledge of subject-specific techniques and methods
	Cc. Knowledge of subject-specific criteria for determining when to use appropriate procedures
D. Metacognitive Knowledge	Da. Strategic knowledge
	Db. Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge
	Dc. Self-knowledge

The cognitive levels hierarchically range with Remembering, Understanding, Analyzing, Evaluating, and Creating. Remembering, the first cognitive dimension, refers to recalling related knowledge from the long-term memory. Following, Understanding refers to determine meaning of instructional messages. Applying refers to the use of procedure learned in a given situation. Analyzing deals with breaking the complex materials into their constituents and connecting them with the principles. Evaluating refers to make judgements based on the standard and criteria. Finally, Creating, the highest level, refers to put together the parts to form a novel and original product. The second dimension, the knowledge dimension, ranges from factual knowledge, conceptual knowledge, procedural knowledge, and metacognitive knowledge. Factual knowledge refers to basic elements related to the field that students have to know to solve problems. Conceptual knowledge concerns with the interconnections among the underlying elements enabling them to work in a tandem. Procedural knowledge refers to ways of things are done based on particularly skills, algorithms, and procedures. Metacognitive knowledge, the highest dimension of knowledge, refers to someone's awareness about his/her thinking.

Table 2. The classification in Bloom's Taxonomy (Krathwohl, 2008)

The Knowledge Dimension	**The Cognitive Process Dimension**					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual	A1	A2	A3	A4	A5	A6
B. Conceptual	B1	B2	B3	B4	B5	B6
C. Procedural	C1	C2	C3	C4	C5	C6
D. Metacognitive	D1	D2	D3	D4	D5	D6

Reading Comprehension Processes

The framework used to the mental process or cognitive dimension is the processes of reading comprehension of Progress in International Reading Literacy Study (PIRLS) reading framework (IEA, 2015). In PIRLS reading framework, the comprehension processes consist of;1) focus and retrieve explicitly stated information, 2) make straightforward inferences, 3) interpret and integrate ideas and information, and 4) evaluate and critique content and textual elements. In addition to comprehension processes, Tengberg (2015) developed the fifth process, that is using knowledge and experience to reflect and evaluate. This is to provide a clear distinction processes between the integrating information and inferencing the textual structure as a text-based activity and global information based on knowledge and experience.

Table 3. Cognitive target descriptors adapted from Tengberg (2015)

Cognitive target	Descriptor
Focus and retrieve	Focus and retrieve explicitly stated information by recognizing and matching appropriate information from a sentence in the text.
Make straightforward inferences	Focus on local meaning by filling gaps at single locations in the text and connecting pieces of information that are relatively clear.
Integrate and interpret	Focus on both local and global meaning by connecting the information that may be implicit or open to interpretation in order to create a more complete understanding.
Examine textual structural and language	Focus on the textual elements of language.
Using knowledge and experience to reflect and evaluate	Focus on local interpretations for which the world knowledge and personal reflections are needed.

Multiple Choice questions to Assess Reading

In the reading questions, there are several ways of testing students' comprehension (Alderson, 2000), such as cloze tests, gap-filling

tests, matching techniques, ordering tasks, short answer tests, etc. However, multiple choice questions (MCQs) are mostly used in reading standardized tests (Campbell, 2005; Khalifa & Weir, 2009; Rowe et al., 2006; Tengberg, 2015). This study investigated reading questions tested in SBMPTN and SIMAK UI by comparing between standard multiple-choice (SMC) and gap-filling multiple-choice (GFMC). In both exams, each correct item is scored +1 (plus 1) and an incorrect item is scored -1 (minus 1).

MCQs are said to be limited to provide higher-order thinking, in terms of test forms that provide test takers problem, information, and algorithms to solve the problems (Moss & Koziol, 1991). However, the idea to construct or manipulate higher-order thinking in MCQs seems to be possible by manipulation of target verbs specific verbs, item flipping, use high quality distractors, tapping 'multiple neuron items'(Morrison & Free, 2001; Scully, 2017) . The manipulation of target specific verbs which is connected to different cognitive processes is done by changing a target verb into its noun derivative and preceding it with a 'knowledge' verb. This will produce stems, for examples; 1) 'select' the best 'description', 2) 'identify' the most accurate 'explanation' (Dickinson, 2011). To this manipulation, though test takers will not generate their own solutions to the scenarios, a 'pseudo-assessment' of higher-order thinking can be produced.

Table 4. Examples of verbs associated with various categories of Bloom's taxonomy (Morrison & Free, 2001)

Verbs Associated with Categories of Cognition for Bloom's Taxonomy of Educational Objectives					
Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
Define	Describe	Apply	Analyze	Compose	Appraise
Identify	Differentiate	Calculate	Categorize	Construct	Assess
Know	Discuss	Classify	Compare	Create	Evaluate
List	Explain	Develop	Contrast	Design	Judge
Name	Rephrase	Examine	Distinguish	Formulate	
Recognize	Restate	Solve	Determine	Modify	
Recall	Reword	Use	Investigate	Plan	

The second is manipulation by item flipping items which are just to remember a concept, the lowest category (Scully, 2017). This, further, can avoid test takers to successfully choose the best option without

understanding the concept. In addition, Dickinson (2011)argues that it can be done by switching the concepts in the items or options and placing the questions of criteria of concepts in the stem. This requires test takers to understand the distractor concepts if they match to the criteria asked in the stem.

Third way is the use high quality distractors. To successfully assess higher-order thinking, test takers are provided the plausible alternatives or distractors (Hancock, 1994; Morrison & Free, 2001). Scully (2017) recommends to provide options that are theoretically plausible to oppose the correct answer, that is 'the best answer'. To do so, it is important to word the item stem appropriately and the item key should be adjusted indisputably and objectively 'more correct' than any other options.

The fourth way to assess higher-order thinking is the use of multiple-neuron items. The 'multiple-neuron' items require the interconnections between knowledge (Scully, 2017). The interconnections, in other words, require the text takers to have more knowledge of more than one fact of concept to finally arrive at the correct answer.

Table 5.The transformation of an item from a 'one-neuron' to a 'five-neuron' adapted from Burns (2010)

1-neuron	Identify *the cell* at the end of the pointer?
2-nueron	Identify *the hormone* produced by *this cell*?
3-neuron	Identify *the target organ* for *the hormone* produced *by this cell*?
4-nueron	Identify the *physiology effect* in the *target organ* for the *hormone* produced by *this cell*?
5-neuron	Identify the *physiology effect* in *the body* caused by *the target* organ for the *hormone* produced *by this cell*?

Scully (2017) then claims that multiple neuron items can be sometimes present in the 'context dependent' in which the stems provide stimulus or scenario that draws test takers' attention to be interpreted based on their knowledge. However, this can be disadvantageous as test takers are bombarded with heavy reading load and high language proficiency (Airasian, 1994). To solve this, test makers can provide information in the forms embedded in the stem of illustration, animation, graph (Dancy & Beichner, 2006).

Methodology

Research Design

The study employs a content analysis (CA) with a quantitative and qualitative design. In this design, CA provides copiable texts to the contexts in which they are used (Baghaei et al., 2020; Krippendorff, 2004). The study adapted the methodology in the study proposed by Baghaei et al. (2020) and Coe and Scacco (2017) that claimed that a quantitative content analysis is concerned to the items which are classified and analyzed. In addition, a quantitative content analysis is aimed at exploring the meanings and themes of messages and how they are organized and presented (Krippendorff 2004). To complete the quantitative content analysis, a qualitative content analysis is employed to investigate, categorize, and analyze the key themes the key themes were investigated, categorized, and analyzed (Baghaei et al., 2020; Bryman, 2004; Krippendorff, 2004).

Materials

The materials used in this study are the Indonesian state university entrance national exams (SBMPTN) and the entrance exams of Universitas Indonesia (SIMAK UI). SBMPTN, for many years, has been used by the state universities in Indonesia as a national entrance test and has been administered by Indonesia's Ministry of research, technology, and higher education. Since 2019 it is now administered by the National Center Test for University Admissions (*LTMPT*, 2021). The SIMAK UI is an entrance test administered independently by Universitas Indonesia (a state university in Indonesia). The materials of both entrance exams are taken from the 2009 to 2018. By selecting a decade period of tests, the researchers attempt to investigate more significant documents.

Table 6. The items of English questions in both the SBMPTN and SIMAK UI 2009-2018

The SBMPTN 2009-2018					The SIMAK UI 2009-2018				
Year	Subjects Tested	Total number of questions	Total number English questions	Time allocation for the whole tests (minutes)	Year	Subjects Tested	Total number of questions	Total number English questions	Time allocation for the whole tests (minutes)

2009	Indonesian Test, Basic Maths Test, English Test	45	15	60	2009	Indonesian Test, Basic Maths Test, English Test	45	15	60
2010					2010				
2011					2011				
2012					2012				
2013					2013				
2014	Academic Potential Test, Indonesian Test	90	15	105	2014		90	15	105
2015					2015				
2016					2016				
2017					2017				
2018					2018				

In terms of number of items, the SBMPTN has more consistent number of questions, to be precisely 15 numbers. However, the SIMAK UI has 20 English questions since 2009 to 2015 and reduced to 15 numbers since 2016.

Coding Schemes and Data Analysis Procedures

As mentioned, the design of this study adapted the study of Baghaei et al. (2020). In their study Baghaei et al., employed a coding scheme based on the Revised Bloom's Taxonomy to codify, classify, and analyze the test questions (see table 2). Furthermore, another coding scheme based on cognitive comprehension processes of IEA (2009) and Tendberg (2019) is also used (see table 3). Finally, to see how MCQs assess higher-order thinking skills, the items are coded based on category; 1) manipulation of target verbs specific verbs, 2) item flipping, 3) use high quality distractors, and 4) tapping 'multiple neuron items'(Morrison & Free, 2001; Scully, 2017). In doing the coding, the researchers codified the data twice in a four-week time interval. When finding the disagreement in the coding, the researchers determined its solution by consulting to the theories and finally decided the final decision after the discussion.

In the process of analyzing data, the researchers used the SPSS 21. To compare the SBMPTN and the SIMAK UI in terms in terms of frequency of lower and higher-order thinking skills, the comprehension processes, and MCQs to promote higher-order thinking, the Fisher-Freeman-Halton Test was performed. To investigate if there were significant differences between the SBMPTN and the SIMAK UI, the chi-square tests were performed.

Findings/Discussion

How Do English Questions of SBMPTN and SIMAK UI Differ in Terms of Types of Questions?

Table 7 depicts the equal distribution of question types tested in the SBMPTN exams since 2009 to 2018 and the dominant questions tested were making inferences in the form of true-false information, clues given in the stems, author's agreement/assumption/implication, prediction/conclusion/hypothesis based on the text. The second dominant questions are the finding specific information which require the test takers to find the specific information in text based on clues in the stem or with WH questions. The following dominant questions are identifying text organization and questions about relation of two paragraphs. This also implies test takers to be familiar with several genres of text and their structures. Several questions about author's purposes, biases, and tones significantly found. Therefore, these questions require comprehension of texts and readers' metacognition of the texts.

Table 7. Types of test items in the SBMPTN and SIMAK UI 2009-2018

Types of test items in SBMPTN 2009-2018	Frequency	(%)	Types of test items in SIMAK UI 2009-2018	Frequency	(%)
Making inference (clues given in the stem/grammatical sentences)	13	8.67	Filling in the gap of a text with a grammatically correct word/phrase/clause	48	25.95
Finding specific information by given clues in the stem (finding a word/phrase/idea in the text)	12	8.00	Filling in the gap of a text with a meaningfully correct word/phrase/clause	37	20.00
Identifying text organization	11	7.33	Completing the dialog/sentence with a grammatically correct expression.	20	10.81
Identifying true/false information/inference	9	6.00	Making inference (clues given in the stem/grammatical sentences)	17	9.19
Identifying the relationship of two paragraphs/texts	8	5.33	Identifying text organization	16	8.65
Asking for topic of a text	7	4.67	Identifying true/false information/inference	11	5.95
Making inference (author's agreement/assumption/implication)	7	4.67	Asking for a word meaning/synonym	8	4.32
Making inference (based on a paragraph/sentences/or text comparison)	7	4.67	Asking for the best title	6	3.24
Asking for a word meaning/synonym	6	4.00	Identifying the author's purpose	5	2.70
Asking for main idea	6	4.00	Identifying the author's tone	4	2.16
Making inference (prediction/conclusion/hypothesis based on text)	6	4.00	Asking for main idea	3	1.62
Deciding the best restatement	6	4.00	Finding specific information by given clues in the stem (finding a word/phrase/idea in the text)	3	1.62
Identifying the author's purpose	6	4.00	Predicting the idea/topic of following paragraph	3	1.62
Identifying the author's tone	6	4.00	Making inference (prediction/conclusion/hypothesis based on text)	2	1.08
Filling in the gap of a text with a meaningfully correct word/phrase/clause	6	4.00	Asking for topic of a text	1	0.54
Identifying word reference	5	3.33			
Making inference (authors' implied meaning based on statement)	5	3.33	Making inference (based on a paragraph/sentences/or comparing 2 texts)	1	0.54

Predicting the idea/topic of preceding paragraph	5	3.33			
Identifying the author's bias	4	2.67			
Predicting the idea/topic of following paragraph	4	2.67			
Identifying the best summary	4	2.67			
Identifying correct analogy based on clues in the text	3	2.00			
Asking for a line	3	2.00			
Creating a relevant question to a text	1	0.67			
Total	150	100.00	185	100.00	

In contrast to SBMPTN, the unequal distribution of question was found in the SIMAK UI exams in 2009 to 2018. The most dominant questions tested are the filling in the gap of a text with a grammatically correct word/phrase/clause. This requires test takers knowledge of English structures and the comprehension of the text. However, test takers might answer the questions just by using their metacognitive of English structure instead of comprehending the texts. The second dominant questions are filling in the gap of a text with a meaningfully correct word/phrase/clause. In this type, test takers must understand the text to be filled in and use their metacognitive to answer the questions. The third dominant types are completing the dialog/sentence with a grammatically correct expression. For this, readers' understanding on spoken language (in the form of dialog and expressions) will significantly contribute to successfully answer the questions. This also seems not to involve students' understanding with a text.

How Do Reading Tests in Both SBMPTN and SIMAK UI Differ in Terms of Comprehension Processes?

By looking at table 8, using the chi-square test (with SPSS 21), it is found that there was significant difference between the SBMPTN and the SIMAK UI in terms of comprehension processes in reading comprehension tests (sig.=.000, p> .05). The finding implies that SBMPTN comprehension processes are distributed more equally than SIMAK UI's. This can be interpreted that SMBPTN comprehension processes range in both processes of integrating and interpreting and processes of examining textual structure and language. However, the SIMAK UI's comprehension processes are dominated by process of examining textual structure and language.

Table 8. Entrance Exams of Comprehension Processes Crosstabulation and Chi-Square Tests of comprehension processes in SBMPTN and SIMAK UI

Entrance Exams * Comprehension Processes Crosstabulation							Chi-Square Tests			
		Comprehension Processes								
		Focus and retrieve	Make straightforward inferences	Integrate and interpret	Examine textual structure and language	Using knowledge and experience to reflect and evaluate		Value	df	Asymp. Sig. (2-sided)
Entrance Exams	SBMPTN	14	28	46	42	20	Pearson Chi-Square	50.171[a]	4	.000
	SIMAK UI	2	8	24	97	13	Likelihood Ratio	52.668	4	.000
Total	16	16	36	70	139	33	Linear-by-Linear Association	25.020	1	.000
							N of Valid Cases	294		
							a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.84.			

How Do English Questions of SBMPTN and SIMAK UI Differ in Terms of Learning Objectives of Revised Bloom's Taxonomy?

Table 9. Bloom's Category in both entrance exams and Chi-Square Tests of comprehension processes in SBMPTN and SIMAK UI of learning levels in the SBMPTN and SIMAK UI exams

Frequency of learning levels presented in SBMPTN and SIMAK UI entrance exams											Chi-Square Tests			
		Bloom Category												
		A1	A2	A3	A4	B2	B3	D2	D3	Total		Value	Df	Asymp. Sig. (2-sided)
Entrance Exams	SBMPTN	9	58	3	71	0	0	6	3	150	Pearson Chi-Square	140.640[a]	7	.000
	SIMAK UI	0	26	0	36	10	65	8	40	185	Likelihood Ratio	179.246	7	.000
Total		9	84	3	107	10	65	14	43	335	Linear-by-Linear Association	97.299	1	.000
											N of Valid Cases	335		

The following objective of the study was to determine if SBMPTN and SIMAK UI were different in terms of learning objectives presented in the Revised Bloom's Taxonomy. In table 9, it shows that there were the absent categories of B2 (conceptual knowledge/understanding cognitive) and B3 (conceptual knowledge/applying cognitive) in SBMPTN exams. In SIMAK UI exams, the absence categories were in A1 (factual knowledge/remembering cognitive) and A3 (factual knowledge/applying cognitive). In SBMPTN exams, the categories of the A3(factual knowledge/applying cognitive) and the A2 (factual knowledge/applying cognitive) were dominantly used. The presence of dominant factual knowledge in SBMPTN implies that most of questions were related to

comprehension of a text. Different from SBMPTN, questions in SIMAK UI were marked by the presence of B3 (conceptual knowledge/applying cognitive), D3 (cognitive knowledge/applying cognitive), and A4 (factual knowledge/analyzing cognitive). This implies that SIMAK UI requires candidates with richer vocabulary, grammars, and comprehension.

Table 9 using the chi-square test (with SPSS 21), shows that there was significant difference between SBMPTN and SIMAK UI in terms of representation of learning levels of Revised Bloom's Taxonomy (sig.=.000, p> .05). This also implies that the distribution of learning levels is not equal in the SBMPTN and SIMAK UI entrance exams.

In relation to higher-order thinking skills, table 10 shows that the equal distribution of higher and lower-order thinking skills is represented in SBMPTN exams. On the hand, lower-order thinking skills are dominantly represented in SIMAK UI exams. This implies the complexity and difficulty of SBMPTN surpasses SIMAK UI. Another implication is that number of questions tested in SIMAK UI were fewer than questions in SBMPTN, that indicates more time consumed in answering SBMPTN questions.

Table 10. Lower-order and higher-order Thinking Skills Crosstabulation and Chi-Square Tests of higher-order and lower-order thinking skills in SBMPTN and SIMAK UI

Lower-order and higher-order Thinking Skills Crosstabulation					**Chi-Square Tests**					
		Thinking Skills		Total		Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
		Lower-order	Higher-order		Pearson Chi-Square	28.330[a]	1	.000		
Entrance Exams	SBMPTN	79	71	150	Continuity Correction[b]	27.092	1	.000		
	SIMAK UI	148	37	185	Likelihood Ratio	28.534	1	.000		
Total		227	227	108	Fisher's Exact Test				.000	.000
					Linear-by-Linear Association	28.245	1	.000		
					N of Valid Cases	335				

Table 10 also reveals that there was significant difference between SBMPTN and SIMAK UI in terms of occurrence of lower and higher-order thinking skills (sig.=.000, p> .05). This also implies SBMPTN exams assessed more critical questions than SIMAK UI exams.

How Do MCQs in SBMPTN and SIMAK UI Promote Higher-order Thinking?

Table 11. MCQ to promote higher-order thinking Lower-order and higher-order Crosstabulation of SBMPTN and SIMAK UI

MCQ to promote higher-order thinking						
		Manipulation of target verbs specific verbs	Use high quality distractors	Tapping 'multiple neuron' items	Using two texts or to paragraphs to compare	Total
Tests	SBMPTN	1	118	13	18	150
	SIMAK UI	0	161	0	0	161
Total		1	279	13	18	311

Table 11 shows that the unequal distribution ways MCQs are developed to assess higher-order thinking in both SBMPTN and SIMAK UI exams. Both exams use high quality distractors in MCQs. However, SBMPTN seemed to be various in using several ways of promoting higher-order thinking, such as the tapping multiple neuron items. Surprisingly, both SBMPTN and SIMAK UI did not employ 'item flipping' as provided in the framework of promoting higher-order thinking in the MCQs. This might happen due to the tests were mostly reading tests, which are different from memorization of concepts, such as in sciences, social, etc. In fact, SBMPTN seemed to apply new ways of promoting higher-order thinking, not mentioned in the framework, that is the providing two texts to compare. This is believed to assess higher-order thinking as the test takers have to compare to find similarities and differences (such as topic, idea and information). This way of MCQs to promote higher-order thinking will require students to experience double or multi-process to arrive at the right answers.

Conclusion

This quantitative-qualitative content analysis was to compare the differences between the entrance exams of the SBMPTN and SIMAK UI in 2009-2018. In terms of types of questions, the SBMPTN provided more equally distributed questions types compared to the SIMAK UI's which are dominated by the questions to fill in the gap of a text with grammatically and meaningfully correct words/phrases/clauses and to complete the dialog/sentence with a grammatically correct expression. The implication of these findings is that secondary high schools need to prepare their students with the relevant vocabulary and grammar knowledge to successfully pass the test. In addition, Universitas Indonesia

(UI) might consider these skills important to be academically successful in the university. However, to the SBMPTN, the implication is that high schools should provide students more various comprehension questions.

In terms of comprehension processes of the two entrance exams, the significant difference is found that the SBMPTN's comprehension processes were more equally distributed and mostly dominated by the processes of integrating and interpreting and the processes of examining textual structure and language. In contrast, SIMAK UI's were dominated by the processes of examining textual structure and language. This implies that the comprehension processes of examining textual structure are both significant in the two entrance exams which mean they are crucial skills needed in the universities. The following implication of the findings is that secondary schools need to facilitate their students with the skills as they are urgently needed in the higher education.

In the distribution of Bloom's learning objectives, both entrance exams are different in which the SBMPTN exams were absent in the categories of the B2 and B3, while the SIMAK UI were absent in the categories of the A1 and A3. The SBMPTN exams were dominated with categories of the A3 and A2 while the SIMAK UI's were dominated by the presence of the B3, D3, and A4. The implication of these findings related to Bloom's category is that the SBMPTN employed questions related to a text comprehension while the SIMAK UI's, beside requiring the comprehension, also employed candidates with wider knowledge of vocabulary and grammar.

In relation to the critical thinking skills, this study presented that the SBMPTN exams promote higher-order thinking skills compared to SIMAK UI. Its implication to these findings related the duration, complexity, and difficulty of the questions is that candidates need to be equipped with the higher-order thinking questions since they were in the high schools. In addition, candidates of the exams need to learn how to answer the questions in a the very limited time.

In terms of how MCQs promote critical thinking, both the entrance exams employed the use of *high-quality distractors*. However, the SBMPTN seemed to apply new ways of promoting higher-order thinking (not mentioned in the framework), that is by providing two texts

to be compared. By this way, to assess higher-order thinking skills, the test takers have to compare by finding similarities and differences (such as topic, idea, and information). This needs double or multi-process to get the right answer. Its practical implication is that candidates of the entrance exams should be prepared with the reading questions *high-quality distractors* and should be introduced with types of reading questions which are to compare two texts.

References

Airasian, P. (1994). *Classroom assessment*. McGraw-Hill.

Alderson, J. C. (2000). *Assessing Reading*. Cambridge University Press. https://doi.org/doi:10.1017/CBO9780511732935

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman.

Baghaei, S., Bagheri, M. S., & Yamini, M. (2020). Analysis of IELTS and TOEFL reading and listening tests in terms of Revised Bloom's Taxonomy. *Cogent Education*, *7*(1). https://doi.org/10.1080/2331186X.2020.1720939

Bloom, B. S. (1956). *Taxonomy of educational objectives*. McKay.

Bryman, A. (2004). *Social Research Methods* (2nd ed.). Oxford University Press.

Burns, E. . (2010). "Anatomizing" reversed: Use of examination questions that foster use of higher order learning skills by students. *Anatomical Science Education*, *3*(6), 330–334.

Campbell, J. R. (2005). Single Instrument, Multiple Measures: Considering the Use of Multiple Item formats to Assess Reading Comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's Reading Comprehension and Assessment* (pp. 347–368). Lawrence Erlbaum.

Coe, K., & Scacco, J. M. (2017). Content Analysis, Quantitative. In J. Matthes (Ed.), *The International Encyclopedia of Communication Research Methods* (pp. 346–356). Wiley-Blackwel.

Dancy, M., & Beichner, R. (2006). Impact of animation on assessment of conceptual understanding in physics. *Physical Review Special Topics*

– *Physics Education Research*, *2*(1), 1–7.

Deygers, B., Van den Branden, K., & Van Gorp, K. (2018). University entrance language tests: A matter of justice. *Language Testing*, *35*(4), 449–476. https://doi.org/10.1177/0265532217706196

Dickinson, M. (2011). *Writing multiple- choice questions for higher-level thinking*. Learning Solutions Magazine. https://learningsolutionsmag.com/articles/804/writing-multiple-choice-questions-for-higher-level-thinking

Hemati, S. J., & Baghaei, P. (2020). A Cognitive Diagnostic Modeling Analysis of the English Reading Comprehension Section of the Iranian National University Entrance Examination. *International Journal of Language Testing*, *10*(1).

IEA. (2015). PIRLS 2016 Assessment framework. In I. V. S. Mullis, M. O. Martin, & M. Sainsbury (Eds.), *PIRLS 2016 Assessment Framework*. Boston College. https://timssandpirls.bc.edu/pirls2016/downloads/P16_Framework_2ndEd.pdf

Khalifa, H., & Weir, C. J. (2009). *Examining Reading: Research and Practice in Assessing Second Language Reading*. Cambridge University Press.

Krathwohl, D. R. (2008). A revision of Bloom's taxonomy : An overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its methodology* (2nd ed.). Sage Publications.

LTMPT. (2021). LTMPT. https://ltmpt.ac.id/?mid=7

Morrison, S., & Free, K. W. (2001). Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education*, *40*(1), 17–24. https://doi.org/10.3928/0148-4834-20010101-06

Moss, P. A., & Koziol, S. M. (1991). Investigating the validity of a locally developed critical thinking test. *Educational Measurement: Issues and Practice*, *10*(3), 17–22.

NACE. (2019). *The four career competencies employers value most*.

https://www.naceweb.org/career-readiness/competencies/the-four-career-competencies-employers-value-most/

P21CS. (2008). 21st Century Skills, Education & Competitiveness. In *Partnership for 21St Century Skills*.

Paribakht, T. S., & Webb, S. (2016). The relationship between academic vocabulary coverage and scores on a standardized English proficiency test. *Journal of English for Academic Purposes*, *21*, 121–132. https://doi.org/10.1016/j.jeap.2015.05.009

PPRC. (2010). 21st century skills for students and teachers. *Kamehameha Schools Research & Evaluation*, 1–25. www.21stcenturyskills.com

Rethinasamy, S., & Chuah, K. M. (2011). The Malaysian university English test (MUET) and its use for placement purposes: A predictive validity study. *Electronic Journal of Foreign Language Teaching*, *8*(2), 234–245. https://doi.org/10.2139/ssrn.2146007

Rowe, M., Ozuru, Y., & McNamara, D. S. (2006). An analysis of a standardized reading ability test: what do questions actually measure? In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the seventh international conference of the learning sciences* (pp. 627–633).

Samad, A. A. (2008). Refining English Language Tests for University Admission: A Malaysian Example. *Asian Journal of University Education*, *4*(1), 57–68.

Scott, C. L. (2015). Education Research and Foresight Working Papers. In *Unesco* (Vol. 1).

Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research and Evaluation*, *22*(4), 1–13. https://doi.org/10.7275/swgt-rj52

Tengberg, M. (2015). National reading tests in Denmark, Norway, and Sweden: A comparison of construct definitions, cognitive targets, and response formats. *Language Testing*, *34*(1), 83–100. https://doi.org/10.1177/0265532215609392