An Investigation into University English Language Instructors' Inclusion of the Revised Bloom's Taxonomy of Cognitive Skills in Testing Language Skills: Selected Universities in Focus

Alemayehu , Wochato, Laiso[1,*], Wondwosen, Tesfamichael, Ali[1] & Zeleke, Arficho, Ayele[2]

[1]English Language and Literature, Wolaita Sodo University, Sodo, Ethiopia

[2]English Language and Literature, Hawassa University, Hawassa, Ethiopia

*Correspdence: English Language and Literature, Wolaita Sodo University, Sodo, P.O. Box 14206 Addis Ababa, Ethiopia. Tel: 251-9-8871-6181

Received: January 4, 2023 Accepted: March 3, 2023 Online Published: March 20, 2023

doi:10.5430/jct.v12n3p1 URL: https://doi.org/10.5430/jct.v12n3p1

Abstract

The purpose of this study was to investigate into university English language instructors' inclusion of the revised Bloom's Taxonomy of Cognitive skills in testing undergraduate students' language skills. The participants of the study were 32 English language instructors who were offering English language courses to undergraduate students. They were selected using a purposive sampling. All the participant-instructors were made to fill in a close-ended questionnaire and a semi-structured interview was held with four of them. Document analysis was also conducted using the judgments of the English language experts. Data collected by the questionnaire was analyzed quantitatively using frequency counts and percentages, whereas data gathered through the interview was analyzed qualitatively. The study employed a descriptive research design and mixed-methods approach. The study's findings revealed that the test items are hardly higher-level cognitive skills (analyzing, evaluating and creating); they are completely dominated by lower-level cognitions (remembering, understanding and applying). Hence, it is recommended that university ELT instructors should give due attention for the revised Bloom's taxonomy of cognitive skills in testing their students' language skills.

Keywords: ELT instructors, inclusion, Bloom's taxonomy, cognitive skills, testing language skills

1. Introduction

Historically, testing began around 1500 years ago in China's imperial system (Angelis et al., 1985). To prevent corruption, Chinese people tested educated people for admission to government jobs for the first time (Ali & Ahmed, 2019; Nejad et al., 2019; Spolsky, 2004). They also administered the National Merit Scholarship Program to high school students for the first time (Angelis et al., 1985; Tournès & Scott-Smith, 2017). It then spread to Western Europe, particularly Italy. It has since evolved in terms of type, purpose, method, approach, and form. As a result, testing in this study refers to tests developed or created by teachers for undergraduates in relation to the revised Bloom's taxonomy of cognitive skill inclusion in a specific educational context and educational context in general.

According to Prasetya (2021), the term "language assessment" is used in free variation with "language testing." However, in the context of this study, the meaning of testing, assessment, and evaluation is used differently. Assessment refers to the process of documenting knowledge, skills, attitudes, and beliefs, usually in measurable terms. The goal of assessment is to make improvements as opposed to simply being judged. In an educational context, assessment is the process of collecting, describing, recording, and interpreting information about learning. Whereas testing measures the level of attitude, skill, or knowledge that has been reached or exceeded. On the other hand, "evaluation" refers to the process of making judgments based on criteria and evidence. As a result, testing in this context refers to the process of measuring language competence(s) in a given context or situation using the revised Bloom's taxonomy of cognitive abilities.

Consider testability when considering English as a foreign language (EFL) testing, specifically whether scores support inferences from test-taker responses to skills, abilities, or knowledge and whether the inferences are generalizable and capable of extrapolation to the real world. Next, simplicity—the requirement that the theory does not use more abstract terms or constructs than are necessary to explain the evidence available Then, coherence—the need to construct theories that are in keeping with what is already known as well as for the theory itself to be internally coherent Finally, comprehensiveness—the requirement that our theories account for as much of the available data and facts as possible—appears.

The cognitive domain focuses on knowledge and the growth of intellectual skills, even though the American Heritage Dictionary of the English Language (Solano-Flores, 2016) defines cognition as "the mental process of knowing, including aspects such as awareness, perception, reasoning, and judgment, or that which comes to be known, as through perception, reasoning, or intuition" (Köksal & Ulum, 2018). The first three levels are lower-order thinking skills, such as remembering, understanding, and applying, while the following three levels are higher-order thinking skills, such as analyzing, evaluating, and creating. The first three levels are lower-order thinking skills, such as remembering, understanding, and applying, while the following three levels are higher-order thinking skills, such as analyzing, evaluating, and creating. As we have already mentioned, this domain is divided into six subsequent thinking levels (Angelis et al., 1985). Each step in Bloom's taxonomy is provided at all levels because it is hierarchical. Consequently, the lower level is also visible from a higher perspective. A person who is performing in the analyzing stage, for example, may also be performing in the remembering, comprehending, and applying stages (Lai & O'Brien, 2020).

There are various circumstances where macro abilities are crucial for language teaching and learning (Kibe et al., 2019). First-language learners and situations for second or foreign-language learning make the problem more challenging (Gerencheal & Mishra, 2019; O'Toole & Junyu Cheng, 2015; Teemant, 2010). Their functional integration is essential, despite the fact that they are taught as different courses in universities (Aydoan & Akbarov, 2014; Tournès & Scott-Smith, 2017). Because they are the foundation of language development, the researcher(s) purposefully chose the four fundamental language abilities. Additionally, it is impressive how the instructors tested these skills in reference to the updated Bloom's taxonomy of educational goals in cognitive skills.

A university is a place where trained, educated, and knowledgeable professionals with the required learning competencies, skills, and understanding are produced. For the process's success, well-equipped teachers in terms of their knowledge, skills, and attitudes are mandatory. Moreover, the instructors' and teachers' instructional approaches and outcomes-measurement practices have great meaning for the institution in general and for the learners in particular. As a result, the inclusion of the revised Bloom's taxonomy of instructional objectives measurement (Chandio, 2021; Muhayimana et al., 2022) caused considerable concern.

Experience shows that there is a gap in universities' English language instructors' testing practices. Hence, the researcher himself had been teaching English as a foreign language for about 18 years, and after more than eight years in a higher education institution, the testing practices suffered lots of problems. For instance, there are experience gaps, inadequate pedagogical training backgrounds, unorganized classroom language, and above all, poor test design. It is possible that it worsens across Ethiopia's universities' from the first to the fourth generation. Moreover, the degree of cruelty varies from generation to generation. That is why the researcher became sensitive to the area and initiated efforts to fill this gap.

Why is there such a chasm? Generally, in Ethiopia, a newly developed curriculum is being implemented, or the policy has changed (MoE, 2021). Second, there is a methodological gap: most of the previous studies used either a quantitative method approach or a qualitative method approach, but not both (a mixed method approach). The current study, on the other hand, employs a mixed-methods approach. Thirdly, the thematic or concept gap means the issues raised in practice are new concepts. Fourthly, the context gap means that most of the studies were conducted at the primary or secondary level; however, this study is based on the context of higher education institutions (universities). Last but not least, the time at which previous studies were conducted was not updated. Therefore, this study generally aims at investigating the university instructors' English language testing practices with regard to the revised Bloom's taxonomy of cognitive abilities for the macro-skills.

1.1 The Purpose of This Study

This study generally aims at investigating the university instructors' English language testing practices with reading to revise Bloom's taxonomy of educational objectives for cognition domain inclusion in undergraduates' macro skills courses. In particular, the inclusion of lower domains of Bloom's taxonomy of cognition and the identification of higher-order cognition test items were investigated.

1.2 Significance of the Study

The researcher believes that there are several areas in which the results of the study could contribute to existing knowledge by providing a detailed analysis of the English language testing practices of the revised Bloom's taxonomy of educational objectives for macro skill courses at tertiary (university) levels. The results also demonstrate instructors' problems that are associated with designing or selecting English language tests at higher education institutions. The results gained may help bring the desired quality to English language testing in general, and they could also bring about changes in learning objectives achievement, test item identification and use, and language skill domain mastery specifically for EFL instructors and teachers. Finally, the findings of the study can be used as a springboard for other researchers who want to research related topics.

1.3 Scope of the Study

The study was conducted at a higher education institution in the South Region of Ethiopia. It included Wolkite and Wachemo (Durame Campus) Universities' English Language and Literature Department instructors. The study was conceptually limited to the testing of English language macro skills and the inclusion of the revised Bloom's taxonomy rather than the broader definition of testing, which includes many other fields of testing.

2. Methods

2.1 Research Design

A descriptive research design was adopted for this investigation. To gather the necessary data, a mixed-method approach was adopted, combining qualitative and quantitative techniques. Because the mixed method approach has a special relevance when we wish to analyzing a problem that is enmeshed in a complex educational and social environment, (Gunasekare, U.L.T.P, 2015; Meakins & Wigglesworth, 2013; Moges, 2018; Rane et al., 2010) the design was chosen over other alternatives.

2.2 Setting, Participants, and Sampling Techniques.

This study was conducted at two National Public Universities in Ethiopia. The participants were instructors from the Department of English Language and Literature (DELLs') who teach undergraduates. As a result, four instructors were interviewed by the researcher using semi-structured interview questions after 32 instructors completed a questionnaire. Purposive sampling was used to choose the teachers for the questionnaires and interviews.

2.3 Methods of Data Analyses

In a mixed methods approach, in many cases, it may be preferable to separate the analyses, only combining the qualitative and quantitative results at the end to illuminate or corroborate each other (Dornyei (2007), i.e., the researcher analyzing d the data in the concurrent (side-by-side) design of the mixed methods approach. In addition, Dornyei (2007), explains that although the analyses were conducted simultaneously, there are times when it may be preferable to keep them separate and only combine the qualitative and quantitative results after the fact in order to clarify or confirm one another. The analyses were done in parallel. According to Creswell (2014, p. 110), this approach helps a researcher achieve what is termed "concurrent triangulation" in mixed methods design, where comparisons between different databases could be made for better effect. According to the same author, this design helps "to actually merge the data and integrate or compare the results of two databases side by side in a discussion." Hence, by adopting this approach, the analysis of data in this study was first made separately for the quantitative and qualitative information, then mixed in the discussion and interpretation section. So, the analyses from both data sets were integrated with answering the research questions of the study (Doyle et al., 2009; Gunasekare, U.L.T.P, 2015; Jang et al., 2008; Johnson & Onwuegbuzie, 2004; Moges, 2018).

2.3.1 Method of Quantitative Data Analysis

Quantitative methods emphasize objective measurements and the statistical analysis of data collected through polls, questionnaires, and surveys, or by manipulating pre-existing statistical data using computational techniques (Rashid & Sipahi, 2021). Quantitative research focuses on gathering numerical data and generalizing it across groups of people or explaining a particular phenomenon. In this study, EFL instructors' responses to Likert's five scale-type questions of the closed-ended questionnaire were categorized into different themes of Bloom's taxonomy of educational objectives in the cognitive domain. Finally, the data were tabulated and analyzing d in the result section using frequencies and percentages (see Tables 1–6) and summarized using minimum, maximum, standard deviation, skewness, and kurtosis (see Table 7).

2.3.2 Method of Qualitative Data Analysis

As stated in the earlier sections, the researcher used two tools for qualitative data collection. These were one-on-one interviews and document (test papers) reviews. The methods for analyzing the data collected by both tools are clearly described in the following section.

2.3.2.1 Analysis of Data via Interview

Since the analysis of data via interview is one of the qualitative data, it used the following procedures: first, all the audio-recorded interviews were transcribed verbatim. Next, transcripts of each university's EFL instructors' interviews were categorized separately. Then, the themes were not totally left to emerge from the data, just like the approach in cognitivists' and constructivists' theories of language testing. Therefore, in the analyses, the transcribed qualitative information was integrated or merged according to the anticipated issues, and it was grouped and organized into categories under major themes. Then, a thematic-based analysis was made. After this process, clear narrations and interpretations were made.

2.3.2.2 Analysis of Data via Document Review

The data that was collected using document review was analyzed based on the themes as stated in the other tools of data collection (questionnaire and interview). This data was sorted based on reviews from secondary data sources. Finally, as stated earlier, at the stage of interpretation, findings from both qualitative and quantitative data were consolidated and discussed in terms of the main strands that describe the EFL instructors' testing practices regarding the inclusion of Bloom's taxonomy of educational objectives for macro skill courses at two National Universities of Ethiopia; then, based on the findings of the study, conclusions were drawn and recommendations were forwarded.

3. Results

As the study focuses on the revised Bloom's taxonomy of educational objectives (remembering, understanding, applying**,** analyzing, evaluating and create) inclusion in EFL test items the quantitative data were analyzing d and found the following results.

Table 1. Summary of "Remembering" Level of Cognitive Domain Related Items

	Scales	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	1	4.11	4.11	**4.13**
	Disagree	4	11.47	11.47	**15.63**
	Agree	13	38.53	38.53	**56.27**
	Strongly Agree	14	46.01	46.01	**100**
	Total	32	100.0	100.0	

3.1 Remembering (lower cognitive skill) it refers that recalling relevant knowledge from long term memory. These questions include: "When I creating EFL tests for macro skills courses, I… ask students to memorize facts that they have learned… elicit learners' definitions of previously learned information… ask the learners to retrieve facts they have previously learned." As a result, is described as follows.

As shown in Table 1, shows the average of three items related to the "remembering" level cognitive skill. So it implies that the majority of the respondents (14, or 46.01%) strongly agreed, and 13 (38.53%) as their test items highly include the lower cognitive "remembering" level domain of revised Bloom's taxonomy educational objectives. It also reveals that the instructors agreed that asking simple questions and terminologies which students memorize them.

3.2 Understanding (lower cognitive skill): it refers to making sense of what you have learnt or cognitive skill that includes interpreting, summarizing and comparing some activity for the given objective of learning. The mean of the items: "When I design EFL test(s) for macro skills course(s), I... give more attention to activities of interpretation, focus on summarizing the language items, and concentrate on comparing language elements." As a result, the mean of the responses were analyzing d and described as follows.

Table 2. Summary of "Understanding" Level of Cognitive Domain Related Items

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Disagree	4	12.4	12.4	**12**
	Undecided	3	9.38	9.38	**22.9**
	Agree	15	46.88	46.88	**70.83**
	Strongly Agree	10	31.25	31.25	**100**
	Total	32	100.0	100.0	

Table 2 shows that the respondent agreed and strongly agreed with the questions. The item types are 15 (46.88%) and 10 (31.25%), which included the majority of the instructors. Only three instructors undecided about their view; the rest, 4 (12.4%), disagreed that they did not include the lower cognitive domain of understanding level. This implies that, still; the majority of the instructors did not test their students in a balanced way. Hence, they give more attention to interpretation, summarization, and comparison than other skills in their test items. Only a few instructors disagreed that they were including different things.

Table 3. Summary of "Applying" level of Cognitive Domain Related Items

	Scales	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	8	25	25	**22.9**
	Disagree	11	34.38	34.38	**56.2**
	Undecided	3	9.38	9.38	**66.66**
	Agree	7	21.88	21.88	**89.57**
	Strongly Agree	3	9.38	9.38	**100.0**
	Total	32	100.0	100.0	

3.3 Applying (lower cognitive skill): it implies that using the knowledge gained in a new way or it is lower cognitive skills that modeling, interviewing and presenting some activity. As we can see, it below the summary of the question and which three items from the questionnaires refer to the level. The questions are: "When I design EFL test(s) for macro skills course(s), I... emphasize more on modeling and simulation activities… deliberately invite students to present different activities… interview students to check their performance." As a result, the mean of the responses were analyzing d and described hereunder.

In the **Table 3** above, the lower cognitive domain of the revised Bloom's taxonomy of educational objectives at the "applying" the third level was tabulated. It showed that the majority of 11 (34.48%) disagreed, and nearly one-quarter of the remaining eight (25% strongly disagreed). Almost a quarter of the respondents (21.88%) agreed, while a few (3, or 9.38%) strongly agreed and three (9.38%) were undecided. This implies that even though there are some instructors who include the third-level "applying" domain questions and items in their tests, the majority of the instructors do not include "applying" level items in their test items. Generally, when we compare to the previous two levels, the percentage of inclusion of the cognitive skills in their test items is decreasing i.e. the higher we go, the lesser the inclusion is.

Table 4. Summary of "Analyzing" Level of Cognitive Domain Related Items

	Scales	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	11	34.38	34.38	**29.7**
	Disagree	8	25	25	**51.55**
	Undecided	2	6.25	6.25	**57.8**
	Agree	6	18.75	18.75	**84.4**
	Strongly Agree	5	15.63	15.63	**100.0**
	Total	32	100.0	100.0	

3.4 Analyzing (higher cognitive skill): It refers breaking the concept into parts and understanding how each part is

related to one another. It is also the higher cognitive domain of revised Bloom's taxonomy of educational objectives. "When I design EFL test(s) for macro skills course(s), I...ask students to differentiate some language functions; …give great emphasis on organizing language function. As a result, the mean of the responses were tabulated as follows.

As shown in **table 4**, the higher cognitive domain of the revised Bloom's taxonomy of educational objectives at the "analyzing" the fourth level was tabulated. It showed that the majority of 11 (34.48%) strongly disagreed, and quarter of the remaining eight (25% disagreed). Less than a quarter of the respondents six (18.75%) agreed, while five (15.63%) strongly agreed and two (6.25%) were undecided. This implies that even though there are some instructors who include the fourth-level "analyzing" domain questions and items in their tests, the majority of the instructors do not include "analyzing" level items in their test items. Generally, when we compare to the previous the inclusiveness of the levels, in test items is highly decreasing i.e. the higher we go, the lesser the inclusion is.

Table 5. Summary of "Evaluating" Level of Cognitive Domain Related Items

	Scales	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	9	28.13	28.13	**26.55**
	Disagree	11	34.38	34.38	**59.4**
	Undecided	2	6.25	6.25	**67.2**
	Agree	8	25	25	**92.2**
	Strongly Agree	2	6.25	6.25	**100.0**
	Total	32	100.0	100.0	

3.5 Evaluating (higher cognitive skill): It refers making judgments based on a set of guidelines. It is the higher level cognitive domain of revised Bloom's taxonomy of educational objectives. It included the questions "When I design EFL test(s) for macro skills course(s), I... request them to make judgments of their learning progress; … frequently request them to report language progress. As a result, the mean of the responses were described and analyzing d as follows.

In **table 5**, the higher cognitive domain of the revised Bloom's taxonomy of educational objectives at the "evaluating ", the fifth level was tabulated. As it has been shown the majority of nine instructors (28.13%) strongly disagreed, and eleven (34.38% disagreed). However, a quarter of the instructors eight (25%) agreed, a few (6.25%) strongly agreed and two (6.25%) were undecided. This implies that even though there are some instructors i.e. about a quarter, who include the "evaluating" level domain questions and items in their tests, the majority of the instructors (in both agree and strongly agree) 62.51% do not include "evaluating " level items in their test items. Generally, when we compare to the previous the inclusiveness of the "evaluating" levels, in test items is still highly decreasing i.e. the higher we go, the lesser the inclusion is.

Table 6. Summary of "Create" Level of Cognitive Domain Related Items

	Scales	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	8	25	25	**25.0**
	Disagree	8	25	25	**50.0**
	Undecided	6	18.75	18.75	**68.75**
	Agree	8	25	25	**93.75**
	Strongly Agree	2	6.25	6.25	**100.0**
	Total	32	100.0	100.0	

3.6 Creating (higher cognitive skill): It refers to putting information together in an innovative way. It is the highest level of cognitive domain of revised Bloom's taxonomy of educational objectives. It included these questions: "When I design EFL test(s) for macro skills course(s), I... always enquire students to rearrange language elements into a new patterns; … ask students to produce language elements into a new structure. The mean of the responses were described next.

As shown in **Table 6**, the highest cognitive domain of the revised Bloom's taxonomy of educational objectives at the "create" level was tabulated. It shows that a quarter or eight (25%) of the instructors strongly disagreed, and quarter

of the remaining eight (25% disagreed). Less than a quarter of the respondents six (18.75%) undecided, while on the other hand eight (25%) agreed and two (6.25%) strongly disagreed. This implies that an average of the instructors still did not include creativity/highest level of cognitive domain enhancing test items in their macro skill tests. One-third (31.25%) of them approved as they include or give chance to enhance the students creativity or putting information together in an innovative way. The rest, six (18.75%), ignored to decide in either disagree or agree which implies uncertainty of the instructors on what they practice in ELT testing.

Table 7. The Lower and the Higher Bloom's Taxonomy Cognitive Skills Comparison

Items Mean	N	Min	Max	Mean	Std. Dev	Skewness		Kurtosis	
Lower Cognitive Skills	32	2.22	4.56	3.5486	.51496	-.291	.414	.543	**.809**
Higher Cognitive Skills	32	1.17	4.50	2.6458	1.12860	.032	.414	-1.636	**.809**
Valid N (listwise)	32								

Generally, as mentioned earlier, the main objective of this study is to identify whether the instructors' test focuses on the higher or lower language skills cognitive domain of the revised Bloom's Taxonomy. So, to identify these issues, 15 questions were asked and nine of them reveal the lower domain language skills and the rest reveal the higher domain language skills.

As shown in **Table 7**, for the minimum value for the lower domains, i.e. remembering the student can recognize and recall relevant knowledge from long-term memory; understanding- the students cam construct meaning from oral, written and graphic messages; applying- the students can use the information in a new way) functional language skills is 2.22 and the maximum value is 4.56. The domain's mean is also 3.55 and negatively skewed. In the same table, for the higher domain (analyzing)-the students can distinguish between parts; how they relate to each other and the overall structure and purpose; evaluating - the student can make judgments and justify decisions; create- the student can put the elements to form the functional whole and to creating a new product or point of view) functional language skills the minimum value is 1.17, and the maximum value is 4.50. The mean for this domain is 2.64 which is skewed positively.

This shows that most of the test items focus on: remembering- the student can recognize and recall relevant knowledge from long-term memory; understanding- the students can construct meaning from oral, written and graphic messages; applying- the students can use the information in a new way) i.e. lower domains functional language skills. It is clear that the EFL tests given by the instructors heavily emphasize less advanced language abilities (lower domain) and fall short of expectations (higher domain). Moreover, the test items hardly include higher domain functional language skills that help the learners analyzing, evaluating and creating language functions.

Regarding to the qualitative data (interview) results are presented as follows. For the same issue from four interviewees requested "To what extent are you clear on the impacts of the revised Bloom's taxonomy language function on EFL testing?" and the respondents reacted to the question as follows :(IA=instructor A, IB=Instructor B, IC=Instructor C and ID=Instructor D)

IA: *Honestly speaking, I have no clear idea about Blooms Taxonomy of language function. Since my background is from applied science. I did not take pedagogy courses and psychology too. The language skills maybe a bit difficult to me to relate with them. IE shares the same idea with IA.*

IB: *I have no updated background for this question.*

IC: *More dominantly lower domain. Sometimes mix it.*

ID: *Okay, the last one (create) is new. In case of understanding regarding with the best level implement or applying undergraduate level and I the end of my undergraduate class there was one course "Measurement and Evaluation" I have experience to prepare the questions when I prepare exam I classify heavy, medium and simple. I try to match with this parting. I rate the simple questions remembering comprehension applying, the most part evaluation. But most of when rating especially at first year level the identify elements and finally weight their work and reason out good or bad paragraph. Why they said give them chance.*

This demonstrates how instructors struggle with a hazy language skills testing area (remembering, understanding, applying, analyzing, evaluating, and create). A few of them are also unaware of the Bloom's taxonomy of language functions, which range from lower domain (remembering, understanding, and applying) to higher domain (analyzing, evaluating, and create). For EFL testing of the four language macro skills for undergraduates, it seems challenging to

construct a test item without a thorough understanding of the cognitive skills of the revised Bloom's taxonomy of language function. Moreover, the document reviews of test items by the discipline's expert judgments revealed the same result as both the questionnaire and interview revealed.

4. Discussion

Despite the variety of measurement tools available, written tests are the most commonly used instrument in academic institutions. The question is how well a component of the test includes the cognitive skills of the revised Bloom's taxonomy of educational objectives (Abu-Rabia & Wattad, 2022). It is also believed that using a systematic balance of questions from both the lower and higher domains of learning in the assessment process, Chandio (2021) claims that the research findings support the scope and function of Bloom's Taxonomy (1956) in improving teaching-learning processes in the classroom. The results of this investigation therefore strongly concur with this notion. Thus, a test questions are essential for assessing students' general cognitive abilities (Berthele & Udry, 2022). Effective test questions should refer to students' varying learning abilities by covering a range of difficulty levels (Annual & Spolsky, 2004; Salmani-nodoushan, 2009). The exposure of students to language discovery improves as the cognitive level in a test increases (Köksal & Ulum, 2018). Because a high-level inquiry necessitates deliberation, evaluation, analysis, creativity, and critical thinking, it may be brand new to the student(Dunn & McCray, 2020). Teachers who create tests to improve their students' higher-order cognitive abilities encourage student interaction (Aydoğan & Akbarov, 2014). As a result, when revising the test questions, they consider how low-level and high-level problems interact with one another. High-level questions are useful for stimulating thought and improving other cognitive skills such as problem solving, decision-making, critical thinking, knowledge-building, and analytical skills whereas low-level questions promote accurate knowledge acquisition and pave the way for the development of high-level cognitive skills (Muhayimana et al., 2022). As a result, the findings of this study clearly demonstrate the practice of English language testing in Ethiopian universities.

5. Conclusions

There are numerous studies that categorize exam questions using the formerly used Bloom's Taxonomy divisions of cognition (Thompson et al., 2008). There haven't been many attempts, nevertheless, teacher-made tests according to the four language skills (reading, writing, speaking, and listening). This study focused on the test questions that particularly addressed four language skills, emphasizing the revised Bloom's taxonomy of cognitive skills inclusion in their EL tests. It also sought to understand instructors' thoughts on how best to incorporate the taxonomy into the testing process as a whole. The matter has been resolved. However, we found that the cognitive levels of remembering and understanding the test questions predominated on the lower-level thinking levels listed in revised Bloom's taxonomy. The application level is included to some extent. Furthermore, the study found that many of the instructors were not even aware of the taxonomy, let alone using it in the higher (analyzing, evaluating and creating) cognitive skills in balanced way. As a result, university ELT instructors are testing their students in undergraduate classes for macro skills courses dominantly simple knowledge that tests the students remembering and understanding of facts and concepts of the English language mastery and highly recommended to re-visit their testing practices regarding the revised Bloom's taxonomy of educational objectives of cognitive skills balanced inclusion in macro skills courses.

Acknowledgements

Our great gratitude goes to Wolkite (WKU) and Wachemo (WCU)-Durame Campus, English Language and Literature Department staff for their invaluable time spent during data collection. Our gratitude goes to my colleagues who encouraged me from many angles.

References

Abu-Rabia, S., & Wattad, H. (2022). Transfer of Language Skills (TLS): A Critical Literature Review. *Creative Education*, *13*(01), 98-111. https://doi.org/10.4236/ce.2022.131007

Ali, I., & Ahmed, K. (2019). A Study of the ESL Students ' Motivational Orientations : A Case of Public and Private Schools. *European Oneline Journal of Natural and Social Sciences, 8*(1), 188-198. Retrieved from http://www.european-science.com

Angelis, P. J., Hughes, A., & Porter, D. (1985). Current Developments in Language Testing. *The Modern Language*

Journal, *69*(3), 293. https://doi.org/10.2307/328354

Annual, S., & Spolsky, B. (2004). *Langauge Policy*. Cambridge University Press, Cambridge. Retrieved from http://www.cambridge.org

Aydoğan, H., & Akbarov, A. A. (2014). The four basic language skills, whole language & intergrated skill approach in mainstream university classrooms in Turkey. *Mediterranean Journal of Social Sciences*, *5*(9), 672-680. https://doi.org/10.5901/mjss.2014.v5n9p672

Berthele, R., & Udry, I. (2022). Multilingual boost vs. cognitive abilities: testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*, *19*(1), 142-161. https://doi.org/10.1080/14790718.2019.1632315

Chandio, M. T., Zafar, N., & Solangi, G. M. (2021). Bloom's Taxonomy: Reforming Pedagogy through Assessment. *Journal of Education and Educational Development, 8*(1), 109-140. http://dx.doi.org/10.22555/joeed.v 8i1.308

Doyle, L., Brady, A. M., & Byrne, G. (2009). An overview of mixed methods research. *Journal of Research in Nursing*, *14*(2), 175-185. https://doi.org/10.1177/1744987108093962

Dunn, K. J., & McCray, G. (2020). The Place of the Bifactor Model in Confirmatory Factor Analysis Investigations Into Construct Dimensionality in Language Testing. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01357

Gerencheal, B., & Mishra, D. (2019). Foreign Languages in Ethiopia: History and Current Status. *Online Submission*, *6*(1), 1431-1439.

Gunasekare, U. L. T. P. (2015). Mixed Research Method as the Third Research Paradigm: A Literature Review. *International Journal of Science and Research (IJSR), 4*(8), 361-367.

Jang, E. E., McDougall, D. E., Pollon, D., Herbert, M., & Russell, P. (2008). Integrative mixed methods data analytic strategies in research on school success in challenging circumstances. *Journal of Mixed Methods Research*, *2*(3), 221-247. https://doi.org/10.1177/1558689808315323

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14-26. https://doi.org/10.3102/0013189X033007014

Kibe, H. M., Owaa, J., Raburu, P. A., Dimaano, M. H., & Hương, N. T. T. (2019). Performance and difficulties in English language macro-skills by freshman students in selected colleges in Bacgiang City, Vietnam. *Journal of English Language and Literature*, *11*(3), 1139-1148. https://doi.org/10.17722/jell.v11i3.416

Köksal, D., & Ulum, G. (2018). Language assessment through Bloom's Taxonomy. *Journal of Language and Linguistic Studies*, *14*(2), 76-88.

Lai, G., & O'Brien, B. A. (2020). Examining Language Switching and Cognitive Control Through the Adaptive Control Hypothesis. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01171

Meakins, F., & Wigglesworth, G. (2013). How much input is enough? Correlating comprehension and child language input in an endangered language. *Journal of Multilingual and Multicultural Development*, *34*(2), 171-188. https://doi.org/10.1080/01434632.2012.733010

MoE. (n.d.). *Newly Developed English Language Curriculum*. [upublished document].

Moges, B. (2018). The implementations and challenges of assessment practices for students' learning in public selected universities, Ethiopia. *Universal Journal of Educational Research*, *6*(12), 2789-2806. https://doi.org/10.13189/ujer.2018.061213

Muhayimana, T., Kwizera, L., & Nyirahabimana, M. R. (2022). Using Bloom's taxonomy to evaluate the cognitive levels of Primary Leaving English Exam questions in Rwandan schools. *Curriculum Perspectives*, *42*(1), 51-63. https://doi.org/10.1007/s41297-021-00156-2

Nejad, A. M., Pakdel, F., & Khansir, A. A. (2019). Interaction between Language Testing Research and Classroom Testing Practice. *Educational Process: International Journal*, *8*(1), 59-71. https://doi.org/10.22521/edupij.2019.81.4

O'Toole, J. M., Junyu Cheng & O'Toole, G. (2015). Language Testing in Asia. In *Language Testing in Asia* (Vol. 5, Issue 1). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=109202997&site=eds-live

Prasetya, R. (2021). English Teaching Based-Strategy LMS Moodle and Google Classroom. *English Education:Journal of English Teaching and Research*, *6*(1), 32-44. https://doi.org/10.29407/jetar.v6i1.15622

Rane, S., He, M., Sayed, D., Yan, L., Vatner, D., & Abdellatif, M. (2010). An antagonism between the AKT and beta-adrenergic signaling pathways mediated through their reciprocal effects on miR-199a-5p. *Cellular Signalling*, *22*(7), 1054-1062. https://doi.org/10.1016/j.cellsig.2010.02.008

Rashid, M. H., & Sipahi, E. (2021). importance of quantitative research in language testing and assessment: in the context of social works. *Linguistics and Culture Review*, *5*(S1), 317-330. https://doi.org/10.21744/lingcure.v5ns1.1413

Salmani-Nodoushan, M. A. L. I. (2009). Research Papers Measurement Theory In Language Testing : Past Traditions By. *3*(2), 1-12.

Solano-Flores, G. (2016). Assessing english language learners: Theory and practice. In *Assessing English Language Learners: Theory and Practice*. Routledge. https://doi.org/10.4324/9780203521953

Spolsky, B. (2004). *Langauge Policy*. Cambridge University Press, Cambridge. Retrieved from http://www.cambridge.org

Teemant, A. (2010). ESL student perspectives on university classroom testing practices. *Journal of the Scholarship of Teaching & Learning*, *10*(3), 89-105. Retrieved from http://files.eric.ed.gov/fulltext/EJ906471.pdf

Tournès, L., & Scott-Smith, G. (2017). Introduction. In *Global Exchanges* (Issue February 2020, pp. 1-30). Berghahn Books. https://doi.org/10.2307/j.ctvw04fqt.6