Developing an effective grammar assessment for junior secondary students: A case study in Batu Layar, West Lombok, Indonesia

Lalu Muhammad Ishak, Baiq Fajriyati
Institute of Teacher Training and education (IKIP) Mataram, lmishak88@yahoo.com

ARTICLE INFO	ABSTRACT
Keywords: *grammar; assessment;* *secondary students*	*The target learners for this assessment are students who are in grade seven and have just entered a junior secondary school in Gerung, Indonesia. Their age is between 12 to 13 years old. The students are a mixed ability group who have a very basic level of English. However, some of them have learnt English in primary school as an optional local content subject. Teachers have identified that most of the students lack the motivation and interest to learn English. As English is taught as a compulsory subject at schools, students have to learn the language and prepare themselves for the national examination. This paper attempts to explore an assessment design for these students which in turn could be used as feedback for both teachers and students in the future. First, the test (assessment) design will be provided in part A This will include the assessment use and timing, questions or tasks, instructions, materials, and the marking (feedback). In part 8, the discussion will address some aspects of language testing including practicality, reliability, validity, authenticity, as well as the positive washback effect.*
Article History:	Submission 12 July 2023 Accepted 19 September 2023 Published 30 October 2023

1. Introduction

The language component to be assessed is grammar, focusing specifically on infinitive verbs and the verb "to be." The assessment design will employ selected responses in which students are asked to perform noticing tasks or consciousness-raising tasks (Brown & Abeywickrama, 2010; Carr, 2011). These models of grammar tests are challenging for junior secondary students, particularly for those learning English as a foreign language. Previous studies have shown that while grammar instruction is crucial for language development, many learners struggle to apply grammatical rules correctly in both written and spoken forms (Gibbons, 2002; Ellis, 2006). Researchers like Schmidt (1990) have emphasized the importance of noticing in language acquisition, suggesting that consciousness-raising tasks can enhance learners' grammatical awareness. However, there remains a gap in understanding how specific grammar features, such as the infinitive verb form and the verb "to be," are best

taught and assessed in secondary school settings. Moreover, there is limited research on how such tasks impact learners' ability to internalize and apply grammar rules over time.

This study aims to address these gaps by exploring the effectiveness of consciousness-raising tasks focused on infinitive verbs and the verb "to be" for junior secondary EFL learners. The following research questions will guide the investigation:

1. How do consciousness-raising tasks impact junior secondary students' understanding and use of infinitive verbs and the verb "to be" in English?
2. To what extent can targeted grammar assessments, such as noticing tasks, help struggling learners improve their grammatical accuracy in English?

2. Literature Review

Grammar instruction and assessment in foreign language learning have been widely explored, with particular emphasis on the effectiveness of various teaching methods. Studies have established that grammar instruction, if integrated appropriately into language learning, can significantly enhance students' overall language proficiency (Ellis, 2006; Swan, 2009). However, despite its importance, many students, particularly those in junior secondary school settings, struggle with grammatical concepts, particularly in English as a foreign language (EFL) context (Gibbons, 2002; Ortega, 2014). One area where learners face considerable difficulty is with the use of infinitive verbs and the verb "to be," which are essential elements of English grammar. These structures are particularly challenging because they have specific syntactic and functional rules that often do not align directly with learners' first languages (Ellis, 2008).

Recent studies on consciousness-raising tasks (CRTs) and noticing activities have shown their potential in enhancing students' grammatical awareness. Schmidt's (1990) Noticing Hypothesis, which posits that learners must consciously notice language forms to internalize them, has been influential in shaping the teaching of grammar. CRTs, which prompt learners to focus on language forms within meaningful contexts, have been found to improve grammatical accuracy and learners' ability to apply rules appropriately in different contexts (Fotos, 1994; Doughty & Varela, 1998). However, while much of the research on consciousness-raising tasks has focused on general grammar instruction, few studies have specifically examined their effectiveness on particular structures like infinitive verbs and the verb "to be" in the secondary school EFL context.

Moreover, existing literature has identified several challenges in designing effective grammar assessments that cater to the needs of secondary school learners. Grammar tests, particularly those that focus on more complex structures like infinitives and auxiliary verbs, are often perceived as daunting by students (Pienemann, 1998). For instance, junior secondary students are still developing cognitive skills necessary for understanding abstract grammatical rules (Tharp & Gallimore, 1988). As such, traditional grammar tests may fail to capture students' actual linguistic capabilities, leading to an underestimation of their potential (Gibbons, 2002). In response, recent calls have been made to design grammar assessments that engage students in tasks that are not only cognitively stimulating but also relevant to real-life language use (Carr, 2011; Gass & Selinker, 2017). This gap suggests a need for more research into how task-based and noticing assessments can help students, particularly in junior secondary education, overcome these hurdles and enhance their grammatical proficiency.

While there has been some investigation into grammar instruction using consciousness-raising tasks (CRTs) and noticing tasks, very few studies focus specifically on their application for teaching infinitive verbs and the verb "to be." Additionally, research on the effectiveness of these tasks in improving secondary school students' grammatical performance, particularly for those who are learning English as a foreign language, remains limited (Tahaineh & Sadeghi, 2013). This gap in the literature underscores the importance of conducting further research to evaluate the impact of noticing tasks on students' grammar, particularly in the context of structures like the infinitive and "to be," which are commonly problematic for EFL learners. Understanding the impact of CRTs on specific grammar points will contribute not only to pedagogical practices but also to the development of more effective assessment tools for grammar teaching.

Given the challenges faced by EFL learners in junior secondary schools and the gaps in the current literature, this study is necessary for a few reasons. First, it will address the lack of research on the effectiveness of consciousness-raising tasks targeting specific grammar forms, such as infinitives and the verb "to be." Second, it will provide insights into how these tasks may help learners overcome common grammar difficulties, especially in an EFL context. Third, it will contribute to the ongoing conversation about the design of more effective grammar assessments that engage students in meaningful learning experiences. Therefore, this study seeks to fill the current research gap by investigating the impact of CRTs on students' mastery of infinitive verbs and the verb "to be," as well as their performance on targeted grammar assessments.

3. Research Methodology

This section outlines the research design, including the participants, methods of data collection, and the approach to data analysis. The purpose of this study is to assess the effectiveness of a diagnostic grammar test focusing on infinitive verbs and the verb "to be," specifically targeting secondary school students learning English as a foreign language (EFL).

Design and Procedure

The study will employ a diagnostic grammar test that is designed to assess students' understanding and correct usage of infinitive verbs (e.g., "to love," "to send," "to live") and the verb "to be" (e.g., "is," "am," "are"). According to Brown and Abeywickrama (2010), diagnostic tests are intended to identify specific aspects of language that students need to improve. In this context, the test will be administered after a review of these grammatical concepts in class, serving as a second round of assessment following initial input sessions.

The test will consist of various tasks aimed at diagnosing students' abilities in using the target grammar forms appropriately. It will include both multiple-choice questions and sentence correction tasks. The tasks will be adapted from the Islamic Life Resource Pack (2010), ensuring that they are suitable for young learners and contextually meaningful (Hasselgreen, 2012). The questions will require students to demonstrate their understanding of the correct usage of infinitives and the verb "to be" within a contextualized text. The text and accompanying tasks will be presented in a way that is engaging and relevant to students, focusing on the practical use of the verbs in everyday scenarios (Wall, 2012).

Participants

The participants in this study will consist of junior secondary school students who are learning English as a foreign language. A total of 30 students, aged 12-15, will be selected from a local secondary school. These students will be at varying levels of proficiency in English, but all will have received prior instruction on the use of infinitive verbs and the verb "to be" in English. The participants will be chosen through convenience sampling, based on their availability and willingness to participate. Parental consent and student assent will be obtained prior to data collection.

Data Collection Methods
The primary method of data collection will be the administration of the diagnostic grammar test. The test will be administered in a classroom setting under controlled conditions. The participants will be given 30 minutes to complete the test, during which they will work individually. A pre-test briefing will be provided to ensure that students understand the instructions and the goals of the test. In addition to the grammar test, students' written responses will be collected for further analysis. This will allow the researcher to evaluate the students' grammatical accuracy and identify specific areas where they are struggling. Teacher observations and feedback will also be recorded to provide additional insights into the students' performance and engagement during the testing process.

Data Analysis
The collected data will be analyzed both quantitatively and qualitatively. Quantitative analysis will involve scoring the students' responses on the grammar test, with particular focus on their accuracy in using infinitive verbs and the verb "to be." Each correct answer will receive one point, and the total score will be calculated for each participant. Descriptive statistics, such as the mean and standard deviation, will be used to summarize the students' performance on the test. Qualitative analysis will involve a detailed examination of the students' incorrect responses. Common errors, such as the use of incorrect verb forms (e.g., "I am like writing letters" instead of "I like writing letters"), will be categorized and analysed to identify patterns of misunderstanding or confusion regarding the target grammar structures. This analysis will also consider any other issues that may arise, such as problems with word order, tense usage, or subject-verb agreement.

Rationale for the Methodology
There are no "magical procedures or formulae" to produce the best test (Bachman & Palmer, 2004, p. 3). The test design and methods outlined above aim to adhere to the principles of classroom language assessment, specifically focusing on validity, reliability, and practicality. The diagnostic test will allow the researcher to identify specific areas where students struggle with infinitive verbs and the verb "to be," which are commonly challenging grammar points for secondary school learners (Gibbons, 2002). By using a contextually relevant and engaging format, the test will not only assess grammatical accuracy but also provide insights into the learning process and areas for improvement.

Ethical Considerations
This study will adhere to ethical guidelines for research involving human participants. Informed consent will be obtained from both the students and their parents, ensuring that

they are aware of the purpose of the study, the procedures involved, and their right to withdraw at any time without penalty. The anonymity and confidentiality of all participants will be maintained throughout the study, and any identifying information will be kept secure and used solely for the purpose of this research.

3. Findings
The test instructions are very important particularly because they inform students 'how they are expected to approach and attempt the test tasks' (Bachman & Palmer, 2004, p. 181). This could be achieved through the following: a) statement of the purposes for which the test is intended, b) language ability that the test is intended to measure, c) specification of the procedures and tasks and d) specification of scoring methods, including criteria for correctness (Bachman & Palmer, 2004).

The application of what Bachman and Palmer (2004) have suggested is outlined in the following example of teacher's instructions given in front of the classroom before the test is attempted. These instructions are deliberately made simple, concise, and where possible teachers use body language to help deliver the message more effectively. A brief version of the instructions is printed on their worksheets, too.

"Hello everyone. Good morning. Today I'm going to give you a grammar quiz which can be done quickly. As I advised you last week, the topic covered in this test is about the appropriate use of 'to be: So, this test is intended to see how well you have understood the subject. Look at this example of the test (the teacher shows an example of worksheet). Here, I've gotten questions which are divided into part A and

B. in part A, you are to identify and correct which 'to be' (is, am, are) is the best answer Underline or circle the only ONE you think the best answer. In part B, all you have to do is to identify the incorrect use of 'to be: Once you have identified the errors, rewrite the full corrected sentences on the space provided on your worksheet.

To save time, you could start working on the easiest question in any part. If you answer them correctly you will be awarded 'excellent: However, if you are only able to answer five or less, your work will be considered 'NTS (not to standard)' (the teacher shows the scoring sheets and grades awarded). Any questions …? Alright, you've got only twenty minutes to do the test; the rest of time will be used to mark them. Now, you can start. Good luck."

4. Discussion
Marking and scoring
Marking can be done by any other proctors including teachers and their colleagues as long as the scoring criteria are clear. In the case of the above test, even students could do the marking (i.e. peer correction/evaluation) by swapping their worksheets providing that the teacher gives them the answer key. Referring to Bloom's taxonomy of grading (Biggs, 1992), all the correct answers will be awarded 'excellent'. Next, having 8 to 9 correct answers will be awarded 'very good'. 6 to 7 correct answers will be 'good' and those getting only 5 or less, will be regarded 'not to standard' (NTS).

Feedback

Letter grades such as A, B, C etc. and numerical scores given to students as marks do not motivate them intrinsically (Brown & Abeywickrama, 201 O). If it is done without comments or feedback, it can reduce the 'linguistic and cognitive performance of students' (p. 39).

Discussion and analysis
In this part of the essay, five components of test design will be discussed: practicality, reliability, validity, authenticity and washback effect.

Practicality
A practical test as suggested by some experts such as Brown and Abeywickrama (2010) and Carr (2011) is one that stays within budget limits, has clear directions for administration, and considers time and effort for both design and scoring. The test designed for Junior Secondary School Gerung was undoubtedly practical for a number of reasons. Firstly, the test can be prepared within a limited budget and only takes a few hours to design. Secondly, the test only takes twenty to thirty minutes to be accomplished altogether. The rest of the time-about twenty minutes-can be used to set up the class at the beginning and in between the sessions. The scoring is even easier as the examiner matches the students' answers with the answer key which takes only one or two minutes. Finally, the test administration does not require extra proctors when the test is conducted. Moreover, only few resources and equipment are needed.

Reliability
The above examples of test items were reliable in a number of ways. Jones (2012) highlighted that a reliable test is consistent in its conditions across two or more administrations; gives clear directions for scoring; has uniform rubrics for scoring; and contains items that are unambiguous to the test taker. The test was considered reliable for a number of reasons. Firstly, regarding learner-related reliability, the test could be conducted through a friendly, casual manner to avoid students getting panicked or too worried. Students missing the test due to sickness or other reasons can sit the test another day as the students 'work is collected by the teacher promptly after the completion of the test.

With regard to rater-reliability, the teacher provides clear scoring criteria which are consistent and applicable to all items. Even if the examiner does the marking in a different situation, the result would be consistent. Regarding the administration reliability, the test can be conducted in a quiet room in the morning where students are still fresh. To avoid students cheating with their neighbors, which is especially common in a large class, the test with different versions (versions A and B) could be distributed to each neighboring student (Carr, 2011). Both versions would contain the same questions with numbers ordered differently. Finally, in terms of test or assessment reliability, it was fairly sure that the test was reliable. Unlike essays which are subjective, the grammar test was considered highly reliable as it is an objective test which has predetermined fixed responses (Bailey, 1998; Brown & Abeywickrama, 201 O; McNamara, 2000).

Validity
Some experts argued that there are no final or exact ways to establish the test validity (Weir, 2005; Broadfoot, 2005; McNamara, 2006 as cited in Brown & Abeywickrama, 2010). However,

a test can be called valid when it measures what is going to measure (Brown & Abeywickrama, 2010). In designing the test, there is no final and absolute measure of validity (McNamara, as cited in Brown & Abeywickrama, 2010).

With regard to content validity, the test was intended to measure students' understanding of the use of verb 'be' in the correct forms. It is obvious that the questions meet the purposes. In terms of criterion validity, the test can be done by observing the use of grammar during English conversations (Brown & Abeywickrama, 201 O). It is predicted that if students were able to complete the tasks successfully, they would be able to appropriately use them in their daily practice.

Regarding the construct validity, as the test requires students to do two types of activities (i.e. identifying mistakes in sentences and selecting one appropriate answer of 'to be'-please see the test sample) it is expected that the test has a construct validity. Whilst in terms of consequential validity, the impact of the test can be seen at a micro level in each individual (Bachman & Palmer as cited in Brown & Abeywickrama, 2010). For example, the test helps students to understand why English sentences do not use verbs with 'to be' (e.g. I am going to school). Finally, the assessment could be perceived as a test by students and therefore has face validity. Even though it is only a case of the 'eye of the beholder' (Brown & Abeywickrama, 2010), this type of validity can increase the students 'trust toward the test.

Authenticity The test was considered authentic because it provides natural language uses for students; the questions are contextualised (e.g. writing emails). It is also meaningful and an interesting topic which is expected to increase the students' motivation: and finally, it was believed that the questions replicated real-world tasks (Brown & Abeywickrama, 2010). In other words, authentic assessment means an assessment which reflects real life language use (Nunan, 1999). In an Indonesian context, the task reflects real life situations in which students often go to internet cafes and make friends online or just have conversations.

5. Conclusion

The test was believed to provide beneficial wash back effect as it positively influences what and how students learn; and gives learners feedback that enhance their language development (Brown & Abeywickrama, 201 O; Wall, 2012). There are a number of cases in which the test above can be claimed to have positive wash back effects. First, the test was done in a direct way, where students do the task without recording or video- taping. The test was also based on the objective of the materials taught at the level. In addition, the test also provides clear feedback which provides praise for good work and constructive criticism for their weaknesses (Brown & Abeywickrama, 201 O; McNamara & Roever, 2006). Finally, as the tasks were understood by the students, they would probably have the motivation to learn more about the subject.

References

Carr, P. (2011). *Grammar for teachers: A guide to American English for native and non-native speakers.* Routledge.

Doughty, C., & Varela, E. (1998). Communicative focus on form. *Forthcoming*.

Ellis, R. (2006). *The study of second language acquisition.* Oxford University Press.

Ellis, R. (2008). The study of second language acquisition (2nd ed.). *Oxford University Press.*

Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28(2), 323-351.

Gass, S., & Selinker, L. (2017). *Second language acquisition: An introductory course* (4th ed.). Routledge.

Gibbons, P. (2002). *Scaffolding language scaffolding learning: Teaching second language learners in the mainstream classroom.* Heinemann.

Ortega, L. (2014). *Understanding second language acquisition.* Routledge.

Pienemann, M. (1998). *Language processing and second language development: Processability theory.* John Benjamins.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158.

Swan, M. (2009). *Practical English usage* (3rd ed.). Oxford University Press.

Tahaineh, Y., & Sadeghi, K. (2013). The role of consciousness-raising tasks in improving grammar knowledge of EFL learners. *English Language Teaching*, 6(4), 50-63.

Tharp, R., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context.* Cambridge University Press.