Evaluating Errors in Exam Papers for Benin Language Programs Credibility

Azoua Mathias Hounnou
University of Parakou, Arafat, Benin

Abstract

Evaluating errors represents an important shortcoming of the teaching-learning evaluation process in every educational system. The purpose of this research was to investigate the causes, consequences, and approaches of solutions to eradicate the evaluating errors in Benin secondary schools. To that purpose, ten (10) teachers and 100 learners were chosen based on their willingness exclusively from GUEMA secondary school in the municipality of Parakou. The data were collected through questionnaires, interviews, group discussions, and analysis of past exam papers. The findings revealed that exam paper errors are often due to the evaluating system that consists of machines. The mistakes concern miswritten words, special notes such as "NB," "Note," etc., ambiguity about items and questions, and unreadability of the paper. Consequently, evaluating errors impacts the learners' present and future language careers. They make learners feel demotivated and stressed, damaging their eagerness to learn the language. It was shown that English as a Foreign Language (EFL) teachers need in-service training, and significant reforms should be done in evaluating the system.

***Keyword*s:** English program credibility, evaluation, errors, exam papers

Introduction

Evaluating errors in testing for program performance constitutes one source of complaint in every educational system. It has been a long-term focus for assessment theorists who seek its minimization and possibly its elimination (Gardner, 2013). It biased the extent of students' learning achievement and performance assessment. A good measurement design aims to minimize the error component. In the simple classical test model, error is thought to occur, but randomly. The importance of random error may be recognized if an assessment is used repeatedly to measure the same individual (Yashim et al., 2021). Then, the observed score would not be the same on each repeated assessment. Scores are more or less variable, depending on the reliability of the assessment instrument (Yashim et al., 2021).

In Benin, evaluation errors still exist, showing the weaknesses of its education programs. Thus, significant changes and reforms need to be made by authorities to improve the effectiveness of the evaluations and the teaching-learning process. These reforms concern the teaching curriculum itself and, most importantly, teachers in class situations.

Teachers' evaluating competence is the master key for measuring students' improvement. Without good evaluation, one cannot appraise whether a learner's teaching-learning process has succeeded and readapt strategies accordingly. From that perspective, a good teacher's quality lies mainly in his ability to design suitable evaluations in accordance with the program taught.

In addition, it is of great importance in the teaching-learning process because it is fundamental for adapting pedagogical responses to students' learning needs (Larrain & Kaiser, 2022).

So, exploring the evaluation system to point out the causes and provide therapy to such everlasting situations that affect the effectiveness of teaching-learning evaluation has become necessary.

Conceptual Clarification

Evaluation is most of the time equated and confused with assessment. In the educational system, assessment determines what a student knows or can do, while evaluation determines the worth or value of a course or program (Iseni, 2011). Whatever the difference, both intend to appraise students' performance of a learned program. In other words, evaluation measures how successful the teaching-learning practice is.

Errors are any bias in a process, and in this context, the gap exists between the result attained and the expected one. In classical test theory, errors are unpredictable, randomly distributed, and occur by chance. By definition, they are expected to have a mean of zero (0) and to be uncorrelated with each other and with all other variables (Yashim et al., 2021)

The technical definition of error can be summarized as any variation from the mark, score, or grade that characterizes the student's 'true' capability in the aspect of their performance that is being assessed (Anikweze, 2010). 'True' in this context generally is the specific level of performance that a student would theoretically achieve if all extraneous error conditions were to be removed, i.e., those not relating to the student's intellectual capability in the assessment task (Yashim et al., 2021)

Thus, the evaluating error is defined as the difference between the distorted information and the undistorted information about

a measured product expressed in its measurements. To sum up, an error is defined as a natural (untrue, wrong, false, no go) value at the output of a measurement system minus an ideal (true, good, right, go) value at the input of a measurement system according to x = xr − xi where x is the error of measurement, xr is the real untrue measurement value, and xi is the ideal actual measurement value (Yu, 2008).

Aim of Study

The present research aim to study the evaluation errors in testing for program performance in secondary schools in Benin country.

Objectives

The main objective of this study was to investigate the leading causes and the impact of evaluating exam errors. Specifically, the study intends to:
1. Analyze the sources of the evaluating errors;
2. Know the impact and interactions extended to learners by considerations to English learning in errors from evaluation paper or test;
3. Find out approaches and solutions in the corrections and the ways to find out or end errors in program performance;

Research Questions

1. What are the impacts of errors in testing programs for performance in the future and immediate affectation against the learners?
2. What are the plausible causes of evaluating errors in the educational system?
3. How can education actors eradicate errors in evaluating students' performance?

Hypotheses

By answering the research questions, some hypotheses are formulated as below:
1. Errors in evaluation are from multiple sources, such as machines, teachers' inefficacy, or system default.
2. The immediate impacts of evaluating errors are the unreliability and invalidity of students' performance, and in the long term, they impact the employees' success at the workplace.
3. To end or minimize the errors in evaluation, one should put important control on machines and the system formatting the evaluation papers.

Methodology

Research Area

The present study was carried out at GUEMA secondary school in the municipality of Parakou. Parakou belongs to the Borgou District in Northern Benin. It is limited to the North by N'Dali municipality and to the South, East, and West by Tchaourou municipality (Hounnou et al., 2022).

Population and Sample

Teachers and students constituted the target population of this research work. One hundred ten (110) respondents were selected, sliced into 100 learners and ten (10) teachers. Both teachers and learners were selected randomly based on their willingness to participate in the survey.

Tools

Methods used for data collection, interviews, questionnaires, group discussion, and last exam papers analysis have been used. The interviews were conducted with English as a Foreign Language (EFL) teachers and aimed to collect information about their experiences and the potential causes of evaluating errors. The forum group discussions with both teachers and teachers helped collect the discourses and cross the responses about the impacts of errors on evaluation sheets on learners. The last exam papers analysis procedure aimed to discover the errors in the papers.

Moreover, questionnaire sheets were distributed to the learners to collect information about their experience with evaluating errors. Questionnaires have been designed in such a way as to facilitate easy understanding by learners.

Data Analysis

Excel 2013 was used as software for data processing. Qualitative and quantitative procedures, consisting of frequency calculation and discourse analyses, were used.

Results

Data from Teachers

The data with teachers reveals the profile of the expert teachers deserving to propose the exam paper. Teachers should meet moral and physical aptitudes before proposing national English exams. Moral aptitudes comprise credibility, the mastership of exam paper methods, a baited CV, expertise in didactic, and completion of tests of aptitudes to the language. The physical aptitudes are a total linguistic mental, good mental health, and a total medical report.

However, 20% of the teachers investigated answered that it is possible to let any teacher or actors propose the exam paper, against 80% who thought the opposite.

Also, the teachers asserted that errors on evaluation sheets are from multiple sources, from teachers, exam authorities, and the system in charge of formatting.

As a consequence, errors in evaluation can cause immediate moral impacts and, thereby, future impacts. For immediate impacts, errors are responsible for disinterest in pursuing the exam, slacking, forsaking of the language, regression in the study, negative perception of the language, and failure. Prospectively, they create hate of the language, drop the level of higher education, and boycott all future perspectives in the language.

According to all the teachers, to put an end to evaluating errors, teachers should be well trained in EFL teaching and student evaluation, and a control system of exam papers must be created. In addition, 70% answered that they can propose exams either for BEPC (Brevet d'Etudes Primaires Certifié) or BAC (Baccalauréat), while 30% said they have no skills to design.

Data from Learners

For the 100% of EFL learners investigated, it's very clear that seeing a mistake on the exam paper is not desirable. During school, 60% have admitted to seeing mistakes on evaluation papers. This negatively affects the understanding of questions or text when there are errors. Since 63% of students struggle to comprehend the objects and questions on the assessment sheets. Also, 60% doubted when seeing errors and 40% could detect mistakes on the sheets. Even so, 57% of students fear telling their teachers about the errors in the proposed exam papers. About the parents implication, 55% of the respondent learners said they had warned their parents about mistakes occurring on their exam papers, either school exams or national exams. During the exam, 97% of the students answered that they are often given free marks and enjoy it, against 3% who dislike free marks. Nevertheless, 70% could correct the mistakes by themselves while 30% could not.

Data from Exam Paper Analysis

The analysis of EFL exam papers for BEPC, BAC, and school exams from 2020 to 2022 has revealed essential shortcomings in evaluating errors. Regarding text quality, 20% of BEPC exams have good-quality text, 67% for BAC, and 20% for school exams. 20% of BEPC exams have fine-quality questions and items quality,

33% for BAC, and 70% for school examinations. Verbs quality is good for 40% of BEPC tests, 100% for BAC, and 50% for school exams. Also, 60% of BEPC exams have good quality adjectives used, 67% for BAC exams, and 70% for school exams. Regarding the readability of sheets, 40% of BEPC exams have good readability quality, 67% of BAC exams are very readable and 90% of school exams are very readable. Moreover, other imperfections like special notes (NB, Notes, and Vocabulary notes) are observed on some exam papers. One can see on BEPC 2010 LV1, BEPC 2017 Normal section, BEPC 2013 LV1, BEPC 2022, Normal Section 2019, BAC 2016 A1 A2 LV2, BAC 2016 B. As concrete examples, in the BEPC 2017 Normal section, it is written, "NB: Tu traiteras les items en anglais" (You will handle the items in English) (Figure 1 to 5). Findings also revealed that mistakes are sometimes related to questions and task formulation. Sometimes, questions are on a table where the answers are not based on (BEPC 2021 Normal Exam) or ambiguous task (BEPC 2020).

Figure 1

BEPC 2019 (presence of additional information "Note")
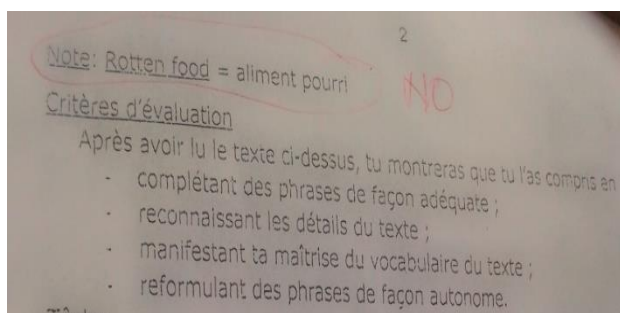
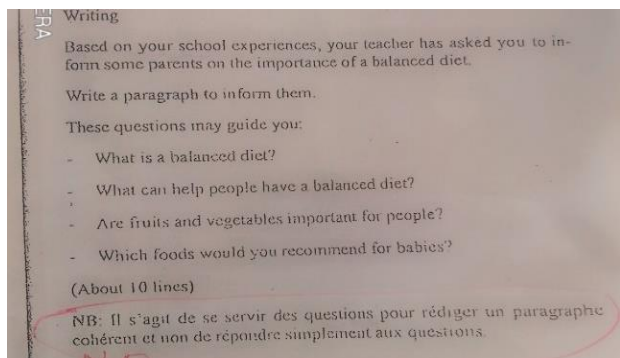*Note.* (Ministry of Secondary School Education, field investigation data, 2023)

Figure 2

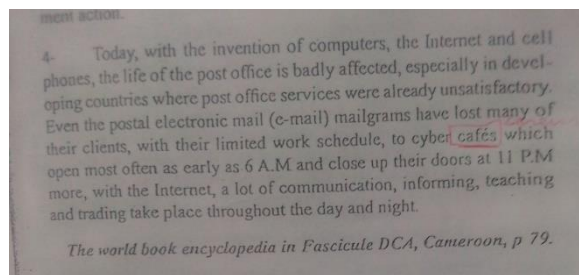BEPC 2021 ('' S '' missing in the word ''Preventions'')

Note. (Ministry of Secondary School Education, field investigation data, 2023)

Figure 3

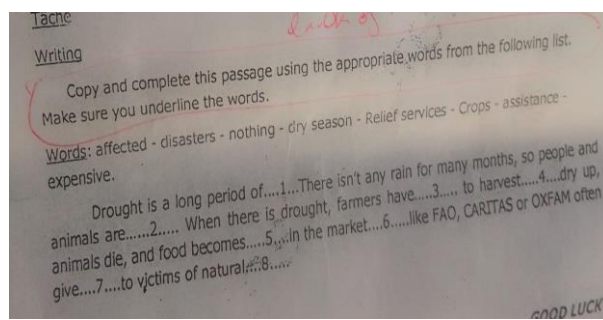BEPC 2016 (presence of special information "NB")

Note. (Ministry of Secondary School Education, field investigation data, 2023)

Figure 4

BEPC 2014 (French word "Cafés" in the text)

Note. (Ministry of Secondary School Education, field investigation data, 2023)

Figure 5

BEPC 2020 (students are asked to copy the text while gaps to fill are numbers "1, 2, 3, 4, 5, 6, 7, 8")

Note. (Ministry of Secondary School Education, field investigation data, 2023)

Discussion

The analysis has revealed that some requirements are needed from EFL teachers before submitting exams for any test, either national or for the school. These requirements lie mainly on the teachers' experiences and expertise. However, 20% have found that anyone can be asked to prepare a test. This has shown the disorganization of the evaluation system, which has no strict rules to be followed. Similar remarks have been made by Assogbadjo et al. (2016). The authors found that it is necessary to rethink the current educational system and implement some strategic actions towards its development.

Dealing with the evaluating errors' sources, the study revealed that many causes can lead to this phenomenon. According to Oates (2017), teacher- and system-based sources are the most important. According to Oates (2017), evaluation systems in all countries are complex. No single aspect of innovation will secure a perfect system because it involves many interacting components such as computers, printers, controlling boards, etc. Suto and Ireland (2021) concluded the opposite when they said that assessment instrument errors stem primarily from human failure.

Shirbagi (2013) goes beyond in his research, pointing out eight sources of bias and error in the perception of students such as errors in instrument content, errors in interpretation of the meaning of ratings, showmanship, errors in instrument reliability, mixed purposes of evaluation, inconsistent methods of instrument administration, common rating error effects, and errors in data implementation.

Those errors lead to mis-measurement or wrong perception of the student's performance, as perceived by Shirbagi (2013). Moreover, the results point out that errors in evaluation can cause immediate moral impacts on the learners and their future. For immediate impacts, errors are responsible for disinterest in pursuing the exam, slacking, forsaking of the language, regression in the study, negative perception of the language, and failure. In the future, they create hate of the language, drop the level of higher education, and boycott all future perspectives in the language.

The study shows that EFL teachers should undergo good training in teaching strategies and evaluation to eradicate evaluation errors. In-service teacher training is a series of activities to engage teachers to enhance their knowledge, raise their skills, and broaden their professional approaches (Ahmed et al., 2021). It is a way to provide continuous education to teachers who have already taken any certification in teaching and have chosen teaching as a profession (Ahmed et al., 2021).

For students, 100% attested to have found errors in the evaluation paper at least once, either for national exams (BEPC or BAC) or school semester exams. They dislike seeing mistakes and, thereby are negatively affected. The majority said they couldn't understand the items and had terrible performance after scoring. 60% doubt when seeing errors and only 40% could detect mistakes on the paper. Nisbet and Shaw (2019) find similar conclusions. As for him, as well as motivating students to pursue their ambitions, fair assessments can build trust and confidence in education systems within society at large.

Also, the findings revealed that 57% of the students fear talking to their teachers due to the teachers' reactions. Similarly, 55% said they had warned their parents about mistakes on their exam papers. The outcomes from the exam papers analysis revealed several shortcomings in the exam papers. Those mistakes vary from miswritten words, special notes such as "NB," "Note," etc., ambiguity about items and questions, and the unreadability of the paper.

Conclusion

The present study has analyzed the evaluating errors in testing for program performance in Benin secondary school. The finding has revealed the sources, the consequences, and some approaches to solutions to overcome this issue. In sum, the research pinpoints the evaluating system as the primary source of mistakes on exam papers. Consequently, it is shown that this phenomenon also impacts the present and future of learners. It makes learners feel demotivated and damper their eagerness to learn English. Thus, essential reforms encompassing teachers' level upgrading through in-service training and improving the evaluating system are needed.

References

Ahmed, H. N., Pasha, A. R., & Malik, M. (2021). The role of teacher training programs in optimizing teacher motivation and professional development skills. Bulletin *of Education and Research*, *43*(2), 17–37.

Anikweze, C. M. (2010). *Measurement and evaluation for teacher education* (2nd ed.). SNAAP Press Ltd.

Assogbadjo, A. E., Idohou, R., & Sinsin, B. (2016). Review of the higher education system in Benin: Status, challenges, opportunities and strategies for improvement. *African Journal of Rural Development*, *1*(2), 139–149.

Gardner, J. (2013). The public understanding of error in educational assessment. *Oxford Review of Education*, *39*(1), 72–92. https://doi.org/10.1080/03054985.2012.760290

Hounnou A. M., Lanmantchion, D. F., Tiebou, L. B., & Adjoha, B. (2022). Students' specific needs of English at Agricultural College. *Djiboul 004* (1), 54–71.

Iseni, A. (2011). Assessment, testing and correcting studentsxs' errors and mistakes. *Language Testing in Asia*, *1*(3), 60–90. https://doi.org/10.1186/2229-0443-1-3-60

Larrain, M., & Kaiser, G. (2022). Interpretation of students' errors as part of the diagnostic competence of pre-service primary school teachers. *Journal für Mathematik-Didaktik*, *43*(1), 39–66. https://doi.org/10.1007/s13138-022-00198-7

Nisbet, I., & Shaw, S. D. (2019). Fair assessment viewed through the lenses of measurement theory. *Assessment in Education: Principles, Policy and Practice*, *26*(5), 612–629. https://doi.org/10.1080/0969594X.2019.1586643

Oates, T. (2017). *A Cambridge approach to improving education*. Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/cambridge-approach-to-improving-education.pdf

Shirbagi, N. (2013). An analysis of the sources of bias and error in students' evaluation of teaching. *Higher Education Letter*, *6*(22), 7-24.

Suto, I., & Ireland, J. (2021). Principles for minimizing errors in examination papers and other educational assessment instruments. *International Journal of Assessment Tools in Education*, *8*(2), 310–325. https://doi.org/10.21449/ijate.897874

Yashim, A. U., Mhab, L. C., & Waziri, J. A. (2021). Measurement errors in educational assessment. *JETP Online Journal*, *1*(1), 1-9.

Yu, F. (2008). Analyst coverage and earnings management. *Journal of Financial Economics*, *88*(2), 245–271. http://doi.org/10.1016/j.jfineco.2007.05.008