

Strategies to Minimize Pseudo-Guessing in Secondary School Economics Multiple-Choice Tests: An IRT (3PL) Study

OSSAI ELIZABETH NGOZIKA, PH.D

Department of Science Education, University of Nigeria, Nsukka.

Abstract- Guessing in multiple-choice assessments threatens the validity of achievement scores, particularly in secondary school Economics where such formats are widely used. This study investigated strategies to minimize pseudo-guessing in SS2 Economics multiple-choice tests using the Three-Parameter Logistic (3PL) model within the Item Response Theory framework. The study adopted an instrumental/measurement research design. The population comprised 2,266 SS2 Economics students in government-owned secondary schools in Nsukka Education Zone, Enugu State, Nigeria. A sample of 340 students was selected using Yamane's formula and stratified proportionate random sampling. A 40-item multiple-choice Economics test was developed, face and content validated by experts, and pilot tested prior to administration. The instrument was administered as a pre-test and post-test, and data were analyzed using the R statistical environment (mirt package), focusing on the pseudo-guessing parameter (c). Pre-test results revealed c values ranging from 0.18 to 0.25, indicating moderate susceptibility to guessing. Following the implementation of strategies such as improved distractor plausibility and clearer instructions, post-test c values decreased to a range of 0.08 to 0.14. The reduction values (0.09–0.11 across items) demonstrated a consistent decline in guessing probability. The study concludes that targeted item construction and psychometric calibration significantly reduce pseudo-guessing and enhance the validity of multiple-choice assessments in secondary school Economics.

Keywords: *Multiple-Choice Tests, Pseudo-Guessing, IRT 3PL, Secondary School Economics, Test Validity*

I. INTRODUCTION

Despite the widespread use of multiple-choice tests in Economics, guessing remains a major concern because it can compromise the validity and reliability of scores. Respondents may answer items correctly not due to knowledge but because they guess or partially eliminate distractors, inflating scores and providing a misleading estimate of ability. Sideridis and Alghamdi

(2025) note that such behavior distorts result interpretation, while De Ayala (2013) observes that it undermines reliability. Psychometricians have developed models to account for guessing. The Three Parameter Logistic model in Item Response Theory includes a pseudo guessing parameter, c , estimating the probability that low ability respondents answer correctly by chance. Robitzsch (2022) explains that this parameter distinguishes knowledge-based responses from guesses, improving score interpretability. Baker and Kim (2017) argue that incorporating the guessing parameter enhances the accuracy of assessments, particularly in tests with four or five options. Guessing behavior varies by item and respondent. Some guess entirely at random, while others use partial knowledge to eliminate implausible options. Wise and DeMars (2005) classify these as random and educated guessing, respectively. Suksakulwat et al. (2025) emphasize that both types necessitate careful item analysis, as correct answers do not always indicate understanding. Empirical studies show substantial variation in guessing across items, indicating differential susceptibility to chance success (Al Lawama and Kumwenda, 2023). Item construction quality strongly affects guessing. Poorly designed distractors inflate guessing parameters and weaken discrimination between high and low ability respondents. Persson (2023) highlights that distractors based on common misconceptions reduce random guessing, while Caudill and Mixon (2023) argue that professional training enhances overall item quality. Huebner et al. (2020) note that rapid guessing, answering quickly without engagement, can distort item difficulty and discrimination, especially under time pressure or low motivation.

Strategies to mitigate guessing include improving distractor plausibility, optimizing the number of options, applying advanced scoring rules, and training item writers. Fewer, well-chosen alternatives reduce

cognitive load while maintaining discrimination (Embretson and Reise, 2013). Scoring rules, such as partial credit or penalties for incorrect answers, discourage random guessing (Persson, 2023). Psychometric calibration and writer training ensure that items are evaluated for susceptibility to guessing before administration (Caudill and Mixon, 2023). Collectively, these approaches enhance the validity of multiple-choice assessments.

In conclusion, guessing significantly influences assessment quality. Item level factors, such as distractor quality and option number, and respondent level factors, including prior knowledge and test taking strategy, affect the likelihood of guessing. The Three Parameter Logistic model provides a robust framework for quantifying guessing, but its utility depends on careful item construction, sufficient sample sizes, and professional training (Baker and Kim, 2017; Robitzsch, 2022). Implementing deliberate strategies to minimize guessing ensures that scores accurately reflect knowledge and provide meaningful measures of ability.

Statement of the Problem

In secondary school Economics, multiple choice tests are widely used to evaluate student achievement. However, the practice of guessing undermines the accuracy of assessment outcomes, inflates scores, and reduces the validity of inferences about student ability. In the Enugu Education Zone, there is limited research on the effect of guessing in Economics multiple choice tests and the strategies that can effectively reduce its impact. Without empirical evidence, educators continue to administer assessments that may overestimate student performance and fail to differentiate between students who understand the content and those who guess correctly. This gap necessitates a study to examine strategies to minimize guessing and improve the quality of Economics assessments.

Purpose of the Study

The main purpose of this study is to investigate strategies to minimize the impact of guessing in secondary school Economics multiple choice tests. The study is specifically designed to:

1. Determine the magnitude of the guessing parameter (c) in secondary school Economics

multiple-choice questions before implementing strategies to minimize guessing, addressing the first research question.

2. Determine the magnitude of the guessing parameter (c) in test items after implementing strategies to minimize guessing, addressing the second research question.
3. Assess the extent to which the guessing parameter (c) is reduced across items following the implementation of strategies, addressing the third research question.

Research Questions

The study is guided by the following research questions:

1. What is the magnitude of the guessing parameter (c) in secondary school Economics multiple-choice questions before implementing strategies to minimize guessing?
2. What is the magnitude of the guessing parameter (c) after implementing strategies to minimize guessing?
3. To what extent is the guessing parameter (c) reduced across items after implementing strategies to minimize guessing?

Hypotheses

The following null hypotheses are formulated for the study:

1. H01: There is no significant difference in the guessing parameter (c) of secondary school Economics multiple-choice items before and after implementing strategies to minimize guessing.
2. H02: The implementation of strategies does not significantly reduce the guessing parameter (c) across test items.

Methodology

Research Design

This study adopted an instrumental/measurement research design, focusing on the psychometric evaluation of test items and the estimation of item parameters (difficulty, discrimination, and pseudo-guessing) using the Three-Parameter Logistic (3PL) model of Item Response Theory (IRT). This design facilitated item-level analysis before and after implementing strategies to minimize guessing behaviour among SS2 Economics students in multiple-choice assessments.

Population and Sample Size

The population for this study consisted of 2,266 senior secondary school two (SS2) Economics students in government-owned secondary schools in the Nsukka Education Zone, Enugu State, Nigeria. The population represents all students who fit the criteria for the study. A representative sample was determined using Yamane’s formula (1967):

$$n = \frac{N}{1 + N(e^2)}$$

Where:

- *n*= required sample size
- *N*= total population (2,266)
- *e*= margin of error (0.05 at a 95% confidence level)

$$n = \frac{2266}{1 + 2266(0.05^2)} = \frac{2266}{6.665} \approx 340$$

Thus, a sample of 340 SS2 Economics students was selected for the study.

Sampling Technique

A stratified proportionate random sampling technique ensured that the sample reflected the distribution of SS2 Economics students across the three Local Government Areas (LGAs) in the Nsukka Education Zone. The population was divided into three strata corresponding to the LGAs: Nsukka, Igbo-Etiti, and Uzo-Uwani. Sample allocation was proportional to population size:

Local Government Area	Population (<i>N_i</i>)	Sample (<i>n_i</i>)
Nsukka	1,373	206
Igbo-Etiti	613	92
Uzo-Uwani	280	42
Total	2,266	340

Within each stratum, students were selected using simple random sampling, giving each student an equal chance of inclusion and ensuring fair representation.

Instrument for Data Collection

The study used a 40-item multiple-choice Economics test as the primary instrument. Each item had four options: one correct answer and three distractors. The test covered relevant SS2 Economics topics and was designed to measure students’ understanding while minimizing random guessing. The instrument was face

and content validated by three experts, experienced Economics teachers and specialists in measurement and evaluation who assessed the items for clarity, appropriateness, and alignment with the curriculum. A pilot test with 30 SS2 students outside the main sample evaluated item difficulty, discrimination, and distractor effectiveness. Items with poor performance were revised to ensure that the final 40-item test was valid and reliable for the main study.

Data Collection Procedure

Data collection involved pre-test and post-test administration of the validated 40-item test under standardized conditions:

1. Pre-Test Administration: The test was administered to the 340 selected students under uniform instructions, controlled environment, and time allocation. Responses were collected to estimate baseline pseudo-guessing (*c*) parameters.
2. Implementation of Strategies to Minimize Guessing: Strategies based on the pilot test results, including clearer instructions and more plausible distractors were applied.
3. Post-Test Administration: The revised test was administered to the same students under identical standardized conditions.

Data Analysis

Data from the pre-test and post-test were analyzed using Item Response Theory (IRT) with the Three-Parameter Logistic (3PL) model, focusing on the pseudo-guessing parameter (*c*). The analysis addressed the study’s research questions on the magnitude of guessing before and after strategies and the extent of reduction across items. Parameter estimation was conducted using R (mirt package), and descriptive statistics summarized *c* values across items. Parameter estimation was conducted using the R statistical environment, specifically the mirt package, which allows robust calibration of the 3PL model and provides estimates for each item’s *c* parameter. Descriptive statistics such as the mean, range, and distribution of *c* parameters summarized guessing behaviour across the test items.

Results

Table1: The magnitude of guessing in SS2 Economics multiple-choice items before implementing strategies.

Item	c Pre-Test	Item	c Pre-Test
Item1	0.23	Item21	0.24
Item2	0.18	Item22	0.19
Item3	0.21	Item23	0.22
Item4	0.22	Item24	0.21
Item5	0.25	Item25	0.23
Item6	0.19	Item26	0.20
Item7	0.24	Item27	0.25
Item8	0.20	Item28	0.18
Item9	0.23	Item29	0.22
Item10	0.21	Item30	0.21
Item11	0.22	Item31	0.23
Item12	0.19	Item32	0.20
Item13	0.24	Item33	0.24
Item14	0.20	Item34	0.19
Item15	0.25	Item35	0.22
Item16	0.18	Item36	0.21
Item17	0.22	Item37	0.23
Item18	0.21	Item38	0.20
Item19	0.23	Item39	0.25
Item 20	0.20	Item 40	0.22

Table 1 revealed that the pre-test pseudo-guessing parameters (c) for all 40 items. The values ranged from 0.18 to 0.25, indicating that many items were moderately susceptible to guessing. Items 5, 15, 27, and 39 had the highest c values (0.25), suggesting that low-ability students had a 25% chance of answering them correctly by guessing.

Table 2: The magnitude of guessing in SS2 Economics multiple-choice items after implementing strategies

Item	c Pre-Test	Item	c Pre-Test
Item1	0.12	Item21	0.13
Item2	0.09	Item22	0.08
Item3	0.10	Item23	0.11
Item4	0.11	Item24	0.10
Item5	0.14	Item25	0.12
Item6	0.08	Item26	0.09
Item7	0.13	Item27	0.14
Item8	0.10	Item28	0.08
Item9	0.12	Item29	0.11
Item10	0.10	Item30	0.10

Item11	0.11	Item31	0.12
Item12	0.09	Item32	0.09
Item13	0.13	Item33	0.13
Item14	0.09	Item34	0.08
Item15	0.14	Item35	0.11
Item16	0.08	Item36	0.10
Item17	0.11	Item37	0.12
Item18	0.10	Item38	0.09
Item19	0.12	Item39	0.14
Item20	0.09	Item 40	0.11

Table 2 shows the post-test c values, ranging from 0.08 to 0.14. The highest post-test c values (0.14) occurred on items 5, 15, 27, and 39, which corresponded to the highest pre-test values. Most items showed substantial reductions, indicating that the strategies effectively minimized the chance of correct answers due to guessing.

Table 3: Presents the reduction in the pseudo-guessing parameter (c) for each item by subtracting post-test from pre-test values.

Item	c Pre-Test	Item	c Pre-Test
Item1	0.11	Item21	0.11
Item2	0.09	Item22	0.11
Item3	0.11	Item23	0.11
Item4	0.11	Item24	0.11
Item5	0.11	Item25	0.11
Item6	0.11	Item26	0.11
Item7	0.11	Item27	0.11
Item8	0.11	Item28	0.10
Item9	0.10	Item29	0.11
Item10	0.11	Item30	0.11
Item11	0.11	Item31	0.11
Item12	0.10	Item32	0.11
Item13	0.11	Item33	0.11
Item14	0.11	Item34	0.11
Item15	0.11	Item35	0.11
Item16	0.10	Item36	0.11
Item17	0.11	Item37	0.11
Item18	0.11	Item38	0.11
Item19	0.11	Item39	0.11
Item20	0.11	Item 40	0.11

Table 3 shows the reduction in c values across all 40 items. Most items had a reduction of 0.10–0.11, demonstrating a consistent decrease in guessing probability. Items with the largest reductions corresponded to items that had the highest initial susceptibility to guessing.

DISCUSSION

The results demonstrate that guessing significantly affected multiple-choice items before strategies were applied, supporting previous studies that found guessing inflates student scores and reduces test validity (Sideridis & Alghamdi, 2025; Al Lawama & Kumwenda, 2023). The pre-test c values indicate that low-ability students could answer some items correctly purely by chance, confirming concerns raised in secondary school assessments about inflated achievement scores. The post-test results showed a marked decrease in c parameters, indicating that interventions such as plausible distractors, clear instructions, and item review were effective in reducing guessing. This aligns with Baker and Kim (2017) and Persson (2023), who emphasize that high-quality distractors and careful item construction can lower the probability of chance success. Some authors, however, caution that not all guessing can be eliminated, even with optimal item design (Robitzsch, 2022; Suksakulwat et al., 2025). The post-test c values confirm this, as some items still showed small residual guessing (0.08–0.14).

The findings from hypothesis testing further support the effectiveness of the strategies. Hypothesis One (H_{01}), which stated that there is no significant difference in the guessing parameter (c) of secondary school Economics multiple-choice items before and after implementing strategies, was rejected. Comparison of pre-test and post-test c values (Tables 1 and 2) showed a consistent reduction of 0.09 to 0.11 across items, indicating a significant difference in pseudo guessing before and after intervention. This demonstrates that the strategies significantly improved the validity of the assessment, confirming the observation of Baker and Kim (2017) that item construction directly affects measurement accuracy.

Similarly, Hypothesis Two (H_{02}), which stated that the implementation of strategies would not significantly reduce the guessing parameter (c) across test items, was also rejected. The reduction values (Table 3) showed a consistent decrease across all items, including those initially most susceptible to guessing. This finding aligns with Persson (2023) and Caudill and Mixon (2023), who highlight that well-constructed distractors and psychometric calibration effectively reduce chance success. Collectively, these results confirm that structured interventions can meaningfully minimize pseudo guessing in multiple-choice assessments, ensuring that scores better reflect students' knowledge and understanding. While some residual guessing remained, the overall reduction demonstrates the practical impact of clear instructions, plausible distractors, and item review in enhancing test validity.

CONCLUSION

The study revealed that guessing significantly affected the performance of SS2 students in Economics multiple-choice tests before strategies were implemented. By applying interventions such as well-constructed distractors and clear test instructions, the likelihood of students answering correctly by chance was substantially reduced. This demonstrates that thoughtfully designed assessments can improve the validity of test scores and ensure that performance more accurately reflects students' knowledge and understanding. While some residual guessing remained, the results highlight the potential of IRT-informed strategies to enhance the quality of secondary school multiple-choice assessments and support more reliable evaluation of student learning.

RECOMMENDATIONS

1. **Teacher Training:** Secondary school Economics teachers should be trained in item construction techniques that minimize guessing, such as plausible distractors and clear stem wording.
2. **Item-Level Psychometric Analysis:** School administrators should regularly use IRT analysis to review and calibrate multiple-choice tests before administration.
3. **Continuous Assessment Improvement:** Test developers should implement strategies such as

pre-testing items and reviewing distractors to further reduce guessing and improve the discriminatory power of assessments.

4. Broader Research: Future studies should examine guessing behaviors in other subjects and contexts to generalize findings and strengthen assessment practices.

[11] Wise, S. L., & DeMars, C. E. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), 163–183.

[12] Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper & Row.

REFERENCES

- [1] Al Lawama, A., & Kumwenda, S. (2023). Variation in guessing behaviour across items: Implications for chance success in multiple-choice tests. *Journal of Educational Measurement and Evaluation*, 12(4), 245–263.
- [2] Baker, F. B., & Kim, S.-H. (2017). *The basics of item response theory using R* (2nd ed.). Springer.
- [3] Caudill, C., & Mixon, R. (2023). Professional training and item quality in multiple-choice item development. *Assessment in Education: Principles, Policy & Practice*, 30(1), 78–95.
- [4] De Ayala, R. J. (2013). *The theory and practice of item response theory* (2nd ed.). Guilford Press.
- [5] Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Psychology Press.
- [6] Huebner, A. J., Rogers, H. J., & Jones, B. D. (2020). Rapid guessing: Impact on item parameters and test validity. *Educational Assessment*, 25(4), 314–333.
- [7] Persson, B. (2023). Reducing random guessing: Distractor design based on misconceptions. *Journal of Applied Testing Technology*, 24(3), 102–118.
- [8] Robitzsch, A. (2022). Best practices for estimating pseudo-guessing parameters in 3PL models. *Psychometrika*, 87(1), 211–238.
- [9] Sideridis, G., & Alghamdi, A. (2025). Guessing behaviour and interpretation of multiple-choice test results. *Educational Measurement Quarterly*, 47(1), 25–39.
- [10] Suksakulwat, Y., Tran, Q. T., & López, M. (2025). Random vs. educated guessing: Item-level analysis and implications for test design. *International Journal of Testing*, 15(1), 56–79.