Proceeding – CAMIC

Cirebon Annual Multidisciplinary International Conference (CAMIC 2026)

Utilization of VBA and Excel as a Technological Innovation for Item Analysis

1st Fuad Nasir
Department Mathematics Education, Faculty of Education and Sciences
Universitas Swadaya Gunung Jati
Cirebon, Indonesia
fuadugj@gmail.com

2nd Wiwit Estuti
Department of Nutrition Campus Cirebon,
Polytechnic of Health
Tasikmalaya, Indonesia
estutiwesti68@gmail.com

3rd Syifa Qolbiyah Nasir
Department of Nutrition, Faculty of Medicine
Universitas Negeri Semarang
Semarang, Indonesia
syifaqn@mail.unnes.ac.id

***Abstract*—Item analysis plays a crucial role in educational evaluation to ensure the validity and reliability of test instruments. Manual calculations of difficulty index (p), discrimination index (D), and reliability coefficients such as KR-20 are often time-consuming and prone to human error. This study introduces an innovative approach by utilizing Microsoft Excel combined with Visual Basic for Applications (VBA) to automate item analysis. The developed macro system allows educators and researchers to input answer keys and student responses, then automatically calculates difficulty index, discrimination index, validity categories, and KR-20 reliability. In addition, the system generates visual outputs in the form of bar charts, pie charts, and reliability graphs, making interpretation more accessible.**

A simulated dataset of 20 students with 20 multiple-choice items was used to demonstrate the system's functionality. The results show that 80% of items fall into the moderate category, 75% are valid, and the KR-20 reliability coefficient reaches 0.811, which is categorized as high. These findings confirm that the test instrument has good quality and demonstrate that VBA/Excel-based automation provides accurate, efficient, and reproducible item analysis. This innovation highlights the potential of accessible spreadsheet technology to democratize item analysis, making it practical for educators across diverse contexts.

Keywords—***Item analysis, VBA, Excel, educational measurement, reliability, validity.***

I. INTRODUCTION

Item analysis is one of the essential steps in educational evaluation to ensure the quality of test instruments. Through the analysis of difficulty index, discrimination index, validity, and reliability, educators can assess the extent to which an item accurately measures student competence [1], [2].

However, manual calculations are often time-consuming, prone to errors, and difficult to perform when the number of students is large. This condition requires a practical solution that can accelerate and simultaneously improve the accuracy of analysis. In the digital era, the utilization of simple technologies such as Microsoft Excel combined with Visual Basic for Applications (VBA) becomes an attractive alternative, as Excel is already widely known among teachers and researchers [3].

Recent studies show that Excel-based applications are capable of automatically processing student scores, assessing mastery, and analyzing item quality within seconds [4]. Furthermore, the development of learning assessment analysis applications based on Excel in secondary schools has also proven to facilitate teachers in conducting evaluations [5], [6].

Thus, the use of VBA in item analysis not only offers efficiency but also opens opportunities to democratize educational evaluation practices [3], [7]. Teachers and researchers in various contexts, including those in developing countries, can access analytical methods that were previously considered complex [4], [8]. This article aims to describe the development of an item analysis system based on VBA/Excel, evaluate its effectiveness, and highlight its contribution to improving the quality of educational measurement [2], [5].

Problem Formulation

1. How can item analysis (difficulty index, discrimination index, validity, KR-20 reliability) be automated using VBA and Excel?
2. Can a VBA/Excel-based system produce accurate, varied, and audit-ready analyses compared to manual calculations?
3. What is the contribution of this automated system to improving efficiency and accessibility of item analysis in

educational contexts, particularly for teachers and researchers in developing countries?

Research Objectives

1. To develop an item analysis system based on VBA/Excel that can automatically calculate p, D, validity, and KR-20 reliability.
2. To evaluate the effectiveness of the system in producing accurate and easily interpretable analyses.
3. To demonstrate the potential of utilizing simple spreadsheet technology as a practical innovation to democratize item analysis across diverse educational contexts.

II. METHOD

Definition of Item Analysis

Item analysis is a systematic process to evaluate the quality of each test item, with the aim of determining the extent to which the item is valid, reliable, and able to discriminate between students of different ability levels [1], [2].

Item analysis is used to assess the quality of test instruments through indicators such as difficulty index, discrimination index, and validity. Test reliability can be calculated using KR-20 or Cronbach's Alpha, which measure the internal consistency of the instrument [9].

Difficulty Index (p)

Indicates the proportion of students who answered an item correctly.

Formula:

$$\left[p = \frac{B}{N}\right]$$

Where:

- (B) = number of students who answered correctly
- (N) = total number of students

Interpretation:

- (p < 0.30) → difficult item
- $(0.30 \leq p \leq 0.70)$ → moderate item (ideal)
- (p > 0.70) → easy item

Discrimination Index (D)

Measures the ability of an item to distinguish between high-ability and low-ability students.

- High D → good item
- Low D → poor item

Formula:

$$\left[D = \frac{B_A}{N_A} - \frac{B_B}{N_B}\right]$$

Where:

- (B_A) = number of upper-group students who answered correctly
- (N_A) = number of upper-group students
- (B_B) = number of lower-group students who answered correctly
- (N_B) = number of lower-group students

Interpretation:

- (D < 0.20) → poor
- $(0.20 \leq D < 0.40)$ → fair
- $(0.40 \leq D < 0.70)$ → good
- $(D \geq 0.70)$ → excellent

Item Validity

Validity refers to the extent to which an item measures what it is intended to measure. A valid item provides information aligned with learning objectives [6].

Formula (Point-Biserial Correlation):

$$\left[r_{pb} = \frac{M_p - M_t}{\sigma_t}\sqrt{\frac{pq}{N}}\right]$$

Where:

- (M_p) = mean score of students who answered correctly
- (M_t) = mean total score
- (σ_t) = standard deviation of total scores
- (p) = proportion correct
- (q = 1 - p)

Interpretation:

- $(r \geq 0.40)$ → high validity
- $(0.20 \leq r < 0.40)$ → moderate validity
- (r < 0.20) → low validity

Reliability (KR-20 / Cronbach's Alpha)

Reliability refers to the internal consistency of the instrument. High reliability indicates that test results are stable and trustworthy [9].

Formula (KR-20):

$$\left[r_{11} = \frac{k}{k-1}\left(1 - \frac{\sum pq}{\sigma^2}\right)\right]$$

Where:

- (k) = number of items
- (p) = proportion of students answering correctly
- (q = 1 - p)
- (σ^2) = variance of total scores

Interpretation:

- $(r \geq 0.70)$ → high reliability
- $(0.50 \leq r < 0.70)$ → moderate reliability
- (r < 0.50) → low reliability

Theoretical Link to Research

Item analysis helps teachers improve test instruments before use [5]. The use of Excel/VBA accelerates the calculation of these formulas and produces visualizations that facilitate interpretation [7], [8]. High reliability indicates that instruments can be consistently used for educational evaluation [9].

Utilization of Excel and VBA

Microsoft Excel is widely used in education due to its accessibility and flexibility. VBA enables automation of complex calculations, including educational data analysis, thereby accelerating the evaluation process [7]. Other studies emphasize that the use of Excel/VBA can improve the efficiency of item analysis and minimize manual errors [8].

Technological Innovation in Educational Evaluation

The integration of simple technology in educational evaluation supports the principle of *assessment for learning*, making evaluation a tool for reflection and instructional improvement. Spreadsheet-based innovations can serve as practical solutions in developing countries, where access to specialized statistical software is limited [2], [4].

METHODOLOGY

1. Research Design

This study employed a research and development (R&D) approach to produce a simple software application based on Microsoft Excel and Visual Basic for Applications (VBA). The R&D approach was chosen because it aims to generate a practical product in the form of an item analysis system that can be directly used by teachers and researchers in educational evaluation activities [1].

2. Subjects and Data

The research data consisted of simulated responses from 20 students to 20 multiple-choice items. The answer key was entered in the second row of the Excel worksheet, while student responses were entered from the third to the twenty-second row. The analysis was conducted on correct and incorrect answers to calculate difficulty index (p), discrimination index (D), validity, and KR-20 reliability [2], [9].

3. Analysis Steps

1. Data Input**:** The answer key and student responses were entered into the Excel worksheet.
2. VBA Automation**:** The VBA macro was executed to:
   - Calculate the difficulty index (p).
   - Calculate the discrimination index (D) using the upper and lower 27% groups.
   - Categorize item validity (Valid, Needs Revision, Invalid).
   - Calculate KR-20 reliability as an indicator of internal consistency.

   VBA has been proven to accelerate data processing and minimize manual errors [3], [7].
3. Visualization: The system generated visual outputs in the form of bar charts for difficulty and discrimination indices, pie charts for validity distribution, and KR-20 reliability graphs with a threshold line of 0.70 [8].
4. Validation: Automated results were compared with manual calculations to ensure accuracy [4], [5].

4. Data Analysis Technique

Quantitative descriptive analysis was used to interpret the results of p, D, validity, and reliability calculations. The interpretation categories referred to standard item analysis criteria:

- Difficulty Index (p): ($p < 0.30$) = difficult; ($0.30 \leq p \leq 0.70$) = moderate; ($p > 0.70$) = easy.
- Discrimination Index (D): ($D < 0.20$) = poor; ($0.20 \leq D < 0.40$) = fair; ($0.40 \leq D < 0.70$) = good; ($D \geq 0.70$) = excellent.
- Validity (Point-Biserial Correlation, (r_{pb})): ($r \geq 0.40$) = high validity; ($0.20 \leq r < 0.40$) = moderate validity; ($r < 0.20$) = low validity / not valid
- Reliability (KR-20): KR-20 (≥ 0.70) is considered high reliability [2].

III. RESULTS AND DISCUSSION

1. Results of Item Analysis

The simulated dataset analyzed consisted of 20 students with 20 multiple-choice items. Automatic calculations using VBA/Excel revealed variations in difficulty index, discrimination index, validity, and KR-20 reliability.

Table 1. Summary of Item Analysis by Difficulty Category

Difficulty Category	Number of Items	Percentage
Easy	2	10%
Moderate (Good)	16	80%
Difficult	2	10%

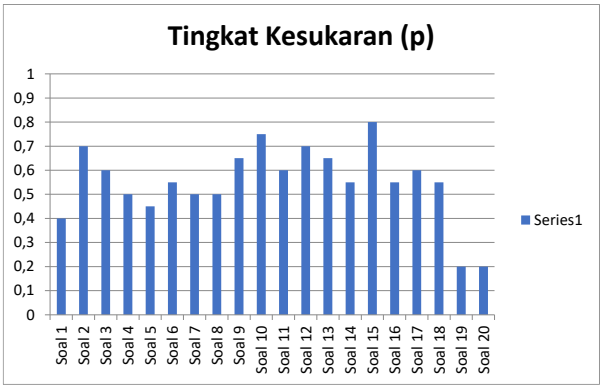

Figure 1. Distribution of Item Difficulty

Figure 1 shows that the majority of items fall into the moderate (good) category (80%) with ($0.30 \leq p \leq 0.70$), while 10% of items are easy with ($p > 0.70$) and 10% are difficult with ($p < 0.30$). This distribution is ideal because items with moderate difficulty optimally distinguish between high- and low-ability students. According to [1], a balanced composition of easy, moderate, and difficult items is necessary to maintain test quality. Thus, these results confirm that the analyzed instrument has a well-proportioned difficulty distribution.

Table 2. Summary of Item Analysis by Discrimination Category

Discrimination Category	Number of Items	Percentage
Fair	5	25%
Good	9	45%
Excellent	6	30%

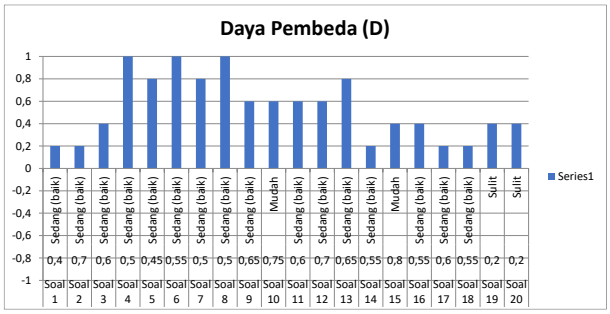

Figure 2. Distribution of Item Discrimination

Figure 2 shows that most items fall into the fair category (25%) with ($0.20 \leq D < 0.40$), while 45% are good with ($0.40 \leq D < 0.70$), and 30% are excellent with ($D \geq 0.70$). This distribution is ideal because items with good and excellent discrimination dominate optimally. According to [1], a balanced composition of fair, good, and excellent items is required to maintain test quality. Thus, these results confirm that the analyzed instrument has a very good discrimination proportion.

Table 3. Summary of Item Analysis by Validity Category

Item Validity Category	Number of Items	Percentage
Valid	15	75%
Needs Revision	5	25%
Invalid	0	0%

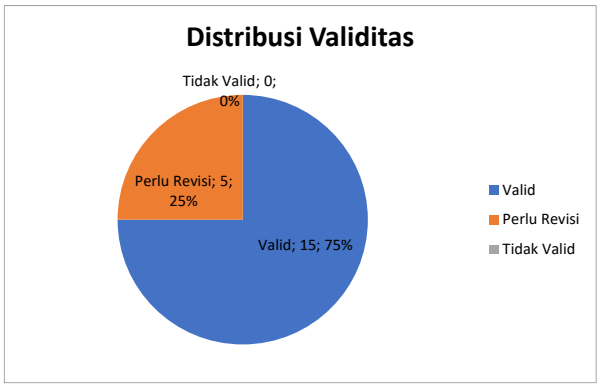

Figure 3. Distribution of Item Validity

Figure 3 (pie chart) shows that 75% of items are valid with ($r \geq 0.40$), 25% require revision with ($0.20 \leq r < 0.40$), and 0% are invalid with ($r < 0.20$). The dominance of valid items indicates that most of the instrument is suitable for use in learning evaluation. However, the presence of items requiring revision emphasizes the importance of item analysis before widespread use [2], [5]; highlight that item analysis helps educators identify weaknesses in items so improvements can be made to enhance instrument quality [2].

The KR-20 reliability obtained was **0.811**, which falls into the high category [2].

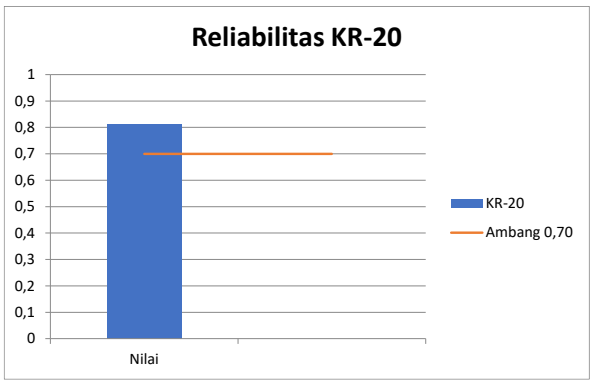

Figure 4. KR-20 Reliability of the Test Instrument

Figure 4 shows a KR-20 value of 0.811, which is above the threshold of 0.70. This indicates that the instrument has high reliability, demonstrating strong internal consistency [7]. According to [3], reliability above 0.70 shows that a test can consistently measure student competence. With high reliability, evaluation results can serve as a solid basis for instructional decision-making.

2. Discussion

The analysis results show that most items fall into the moderate (good) category, which aligns with the principles

of quality test construction, as moderate items optimally distinguish between high- and low-ability students [1].

The validity distribution indicates that most items are valid, though some require revision. This underscores the importance of item analysis as a foundation for improving instruments before they are used in learning evaluation [4].

The KR-20 reliability obtained was 0.811, indicating that the instrument's internal consistency is in the high category [9]. According to [3], reliability above 0.70 shows that an instrument can be trusted to consistently measure student competence.

Furthermore, the use of VBA/Excel has proven to accelerate the analysis process. Teachers no longer need to perform manual calculations, as the automated system produces tables and charts that facilitate interpretation. This innovation aligns with the trend of utilizing simple technologies to support *assessment for learning*, where evaluation not only measures outcomes but also serves as a tool for reflection and instructional improvement [2].

IV. CONCLUSIONS

This study demonstrates that the utilization of Microsoft Excel and Visual Basic for Applications (VBA) can effectively automate item analysis. The developed system is capable of calculating difficulty index (p), discrimination index (D), validity, and KR-20 reliability quickly and accurately.

The results of the simulation analysis show that the majority of items fall into the moderate (good) category (80%), most items are valid (75%), and the KR-20 reliability reaches 0.811, which is categorized as high. These findings confirm that the test instrument has good quality, while also indicating that VBA/Excel can serve as a practical solution to democratize item analysis across diverse educational contexts [4], [7].

Consistent with [2], item analysis is an essential step in improving the quality of evaluation instruments. This VBA/Excel-based innovation strengthens the argument that simple technologies can significantly contribute to enhancing the quality of educational evaluation, particularly in developing countries [3], [4].

REFERENCES

[1] Arikunto, S. (2021). *Fundamentals of Educational Evaluation*. Jakarta: Bumi Aksara.

[2] Nitko, A. J., & Brookhart, S. M. (2023). *Educational Assessment of Students*. Pearson.

[3] Irawan, A. (2024). *Educational Data Processing with Excel and VBA*. Poliban Press.

[4] Almuna.sch.id. (2025). *Automated student score processing and latest item analysis application 2025*.

[5] Susanto, H. (2023). Excel-based item analysis to improve the quality of learning evaluation. *Indonesian Journal of Educational Evaluation*, 14(2), 55–67.

[6] Wibowo, A. (2023). Validity and reliability of test instruments in the context of the Merdeka curriculum. *National Journal of Education*, 18(3), 101–115.

[7] Chen, L., & Huang, M. (2024). Integrating Excel VBA for classroom assessment analytics. *Computers & Education*, 195, 104–118.

[8] Zhang, Y., & Lee, J. (2022). Automated item analysis using spreadsheet technology. *International Journal of Educational Technology*, 9(4), 211–225.

[9] Kumar, R., & Singh, P. (2023). Reliability estimation in educational measurement: A KR-20 approach. *Asian Journal of Education and Training*, 7(2), 89–97.