Edureligia : Jurnal Pendidikan Agama Islam 09 (02) (2025) : 223-239
DOI: http://dx.doi.org/10.33650/edureligia.v9i2.10366

Edureligia
Jurnal Pendidikan Agama Islam
https://ejournal.unuja.ac.id/index.php/edureligia
p-ISSN : 2549-4821
e-ISSN : 2579-5694

DEVELOPMENT OF A MULTIPLE-CHOICE ITEM ANALYSIS APPLICATION TO ENHANCE LEARNING ASSESSMENT INSTRUMENTS

Siti Citra Aulia Aprina[1], Hilyah Ashoumi[2](✉), Muhammad Kris Yuan Hidayatulloh [3], Munifah[4], Iskandar Tsani[5]
[1,2] Universitas KH. A. Wahab Hasbullah, Jombang, East Java, Indonesia
[3]Universitas Negeri Surabaya, Surabaya, East Java, Indonesia
[4,5]Universitas Islam Negeri Syekh Wasil Kediri, East Java, Indonesia

Article History:
Received: April 2025
Accepted: October 2025
Published: December 2025

Keywords:
Developent Assessment Instruments, Question Item Analysis, Multiple Choice

(✉)**Correspondence to:**
hira@unwaha.ac.id

Abstract: The research aimed to ascertain the viability, receptivity of teachers to the implementation of multiple-choice item analysis and to develop a multiple-choice item analysis application in desktop form. This type of research was developed using the research and development (R&D) method. This research employs the ADDIE development model, which comprised five stages. The phases of analysis, design, development, implementation, and evaluation were employed. The results of this research indicated that the level of feasibility, based on the first media expert with a percentage value of 85%, was declared "Very Eligible"; the second media expert, with a percentage value of 86.66%, was declared "Very Eligible"; and the third media expert, with a percentage value of 95%, was declared "Very Appropriate." In contrast, the response results from teachers received a percentage value of 80.5%. It indicated the application was suitable for use in educational institutions. The findings showed that the multiple-choice item analysis application had great potential to improve the quality of test questions in schools. The application overcame the weaknesses of existing software by offering a reliable, easy-to-use tool, supported by high ratings from experts and positive feedback from teachers. This innovation helped educators create better assessments, improving learning outcomes and overall education quality. The implication of this research was that the developed application can serve as a practical solution for schools to strengthen assessment quality, enhance teachers' competency in test construction, and support policy-making in education by providing accurate data-driven insights into student learning outcomes.

Please cite this article in APA style as:
Aprina, S. C. A., Ashoumi, H., Hidayatulloh, M. K. Y, Munifah, M., & Tsani, I. (2025). Development of A Multiple-Choice Item Analysis Application to Enhance Learning Assessment Instruments. *Edureligia: Jurnal Pendidikan Agama Islam*, *9*(2), 223-239.

INTRODUCTION

The advent of information and communication technologies (ICT) has revolutionized numerous aspects of modern life, and education stands as one of the fields most profoundly impacted. Educators worldwide are increasingly adopting innovative approaches to enhance teaching and learning processes, recognizing the potential of these technologies to engage students and improve educational outcomes. Among the most significant contributions of ICT is the Internet, which has fundamentally transformed how information is accessed, shared, and utilized. The integration of the Internet into educational practices has enabled the development of dynamic, interactive, and personalized learning environments that cater to diverse learner needs (Halverson et al., 2020). Furthermore, communication technologies have reshaped the concept of basic needs, embedding themselves into the fabric of daily life and influencing human interaction, work, and knowledge acquisition. In education, the availability of these technologies has shifted the paradigm from teacher-centered to learner-centered approaches, emphasizing collaboration, critical thinking, and digital literacy (Rozi et al., 2022).

Tools such as virtual classrooms, online libraries, and digital assessment platforms empower students to learn at their own pace while giving educators the means to track progress and provide timely feedback. This technological integration also facilitates access to quality education in remote and underserved areas, bridging gaps in educational equity (Astiza et al., 2023). As ICT continues to evolve, its role in education expands, offering unprecedented opportunities to redefine learning experiences and address challenges such as engagement, resource scarcity, and skill alignment with the demands of the 21st century workforce (Laato et al., 2023).

The integration of information technology, particularly the Internet, in education has transformed the way students access and engage with knowledge, offering unparalleled convenience and expanding learning opportunities (Belkhir, 2024). This technological shift allows students to obtain learning materials anytime and anywhere, enhancing their educational experience and fostering independent learning. Moreover, it enables teachers to deliver lessons more effectively by utilizing digital resources such as multimedia presentations, online assessments, and interactive tools, which cater to diverse learning styles (Chin et al., 2022). By incorporating technology into the classroom, teachers can create a more engaging and dynamic learning environment that encourages student participation and fosters enthusiasm. Technology also supports the personalization of learning, allowing for adaptive learning pathways and the tailoring of content to individual student needs (Mai et al., 2024). In achieving educational objectives, however, the role of structured teaching procedures remains crucial. These procedures are not only institutional guidelines but also serve as strategic frameworks that ensure the coherence and effectiveness of the teaching process.

It is within these frameworks that the measurement of student understanding and learning success becamed essential. Implementing

evaluations was necessary to assess how well students have grasped the material and to identify areas that require further attention (Forsblom et al., 2022). However, the success of evaluations depended significantly on how the learning outcomes were analyzed. To optimize the evaluation process, a thorough analysis of student performance was required to ensure that feedback is both actionable and meaningful. This analysis provided valuable insights into the effectiveness of teaching methods and highlights areas for improvement in both teaching practices and curriculum design. Ultimately, a well-structured approach to teaching, combined with effective evaluation, fosters an environment in which both students and educators can continuously grow and achieve academic success.

Evaluation is the final process of learning. Evaluating learning outcomes involves measuring students' knowledge and skills after they have received learning materials from the teacher (Peters, 2022). One key instrument for measuring learning progress and understanding students' academic improvement is the administration of tests provided by the teacher. Tests serve as benchmarks for determining student learning success and are expected to provide information that reflects the actual learning context (You et al., 2021). Test instruments are considered more effective in assessing students' understanding of the material. Additionally, tests help determine the level of a learner's ability (Ribeiro-Silva et al., 2022).

Given the critical role of evaluation in the learning process, establishing a robust evaluation system is paramount. Many teachers still face challenges in assessing learning outcomes, particularly in defining evaluation criteria, designing assessment tools, ensuring item quality, and establishing scoring guidelines (Slepkov et al., 2021). To achieve reliable and valid assessment results, it is essential to use instruments that accurately measure targeted objectives and consistently yield reproducible data. A comprehensive assessment system should embody characteristics of validity, reliability, relevance, practicality, specificity, distinctiveness, and proportionality (Ismail, 2020). Such a system not only provides accurate insights into students' abilities but also serves as a benchmark for gauging learning outcomes, thereby encouraging effective teaching strategies and fostering students' enthusiasm for learning (Darvin & Norton, 2023).

This research focused on developing assessment instruments aimed at evaluating student learning outcomes to assess their comprehension and critical thinking abilities. The instruments utilized were multiple-choice tests featuring four answer options, where students select the correct response for each item. Multiple-choice testing was characterized by its objectivity, ensuring clear and unambiguous answers that facilitate impartial evaluation (Hidayatulloh & Ashoumi, 2022). Each question within these tests required learners to select the correct answer from the provided options, with predetermined scoring rules ensuring consistent evaluation across all responses. This systematic approach guarantees uniformity in scoring, thereby maintaining the reliability and validity of the assessment process (W. Liu, 2021).

Based on interviews conducted with teachers at MTs. Umar Zahid Semelo, it was found that teachers face challenges in evaluating learning. They encounter difficulties in developing high-quality assessment instruments and struggle with compiling assessment grids. These difficulties stem from limited access to software for item analysis, which complicated the process of creating effective assessment questions. The consequence of not analyzing these items was that students may not engage optimally with the questions, thus rendering them ineffective as benchmarks to assess students' comprehension and learning success.

One way teachers attempt to assess the quality of learning outcome assessment instruments is through item analysis (Rezigalla et al., 2024). High-quality test instruments are characterized by high validity and reliability (Belay et al., 2022). In this study, the item analysis application utilized was used the Anabus application, a desktop-based tool designed for analyzing each item of multiple-choice questions, descriptions, and questionnaires. This application's functionality included determining weighted scores, assessing question reliability, calculating correlations between item scores and total scores, establishing scoring and superior groups, and evaluating difficulty levels and discriminant power (Vaughn et al., 2023).

The findings from prior studies indicate that utilizing the Object Analytical Data Processing (OASIS) application, coupled with optimizing Microsoft Excel 2013 features, facilitates the analysis of multiple-choice question performance (Khoirunnisa, 2020). This aligned with the objective of developing an item analysis application tailored to diverse contexts and topics as identified in earlier research. The aim of developing this application was to simplify the process for teachers to assess test instruments, thereby enabling them to evaluate the suitability of these instruments before implementing them in student assessments.

This study aimed to determine the feasibility of multiple-choice item analysis applications, to determine the teacher's response to multiple-choice item analysis applications, and to develop multiple-choice item analysis applications in desktop form. The development of this application was to facilitate teachers in the process of preparing quality question instruments. Therefore, the author was interested in examining the feasibility and teacher responsed to the multiple-choice item analysis application to improve the quality of learning outcomes instruments.

The importance of this research lied in addressing the challenge teachers face in developing and evaluating quality multiple-choice assessment instruments due to limited access to item analysis software. Unlike previous studies focused mainly on optimizing existing tools like Microsoft Excel or generic item analysis applications, this research develops a dedicated desktop-based Anabus application tailored specifically for educational practitioners. The application combined sophisticated statistical analysis with user-friendly features that improve educators' ability to evaluate and refine test items effectively. The author was actively involved in designing and developing

Anabus, ensuring that the tool meets actual classroom assessment needs and incorporates feedback from teachers. This made the study interesting as it bridges the gap between complex statistical methods and practical usability in diverse educational contexts.

RESEARCH METHOD

The research method used in this study was the Research and Development (R&D) method. The product developed in this research and development included software for learning, guidance, training, evaluation, and other related aspects. The resulting application from this development was a software-based item analysis application, specifically the Anabus application. This study followed the Research and Development (R&D) approach, which involves enhancing existing software (Mesra, 2023). R&D comprised a series of processes aimed at developing new products or improving existing ones to meet specific criteria. This research adhered to the ADDIE development model, which encompasses five stages: Analysis, Design, Development, Implementation, and Evaluation (Kumar et al., 2021).

The research followed the five stages of the ADDIE development model. Beginning with the Analysis stage, data collection informed the development of the Anabus application, designed for analyzing multiple-choice items. Moving to the Design stage, researchers focus on creating the application's blueprint and establishing criteria for question validity and reliability. In the Development stage, the application toook shape as per the design specifications, marking the start of testing. Subsequently, the Implementation stage tests the application's functionality and appearance, starting with evaluation by media experts and later by teachers. Finally, the evaluation stage concluded the research, assessing the application's effectiveness through validation by experts and trials in educational settings.

RESULT AND DISCUSSION

Result

The product developed by the author is an application designed to enhance the quality of student learning outcomes through the analysis of multiple-choice items. Developed using the ADDIE model, the design process adheres to the stages inherent in this model. These stages in developing the multiple-choice item analysis application include: Analysis, where information gathering informs product development; Design, focusing on the application's structural blueprint; Development, translating the design into a functional product; Implementation, testing the application's functionality and usability; and Evaluation, assessing the application's efficacy through expert validation and teacher trials (Balasopoulou et al., 2017).

Design and Analyze Stage

The current stage focused on analyzing the concept of a product, particularly emphasizing improvements to address operational limitations observed in its previous version. The prior product had minimal features and

operational capabilities, which posed challenges for users in efficiently utilizing it. Therefore, the development process aimed to enhance the product's functionality, making it significantly easier and more user-friendly. This improvement was driven by the need to ensure better usability, efficiency, and satisfaction for the users. Additionally, the development emphasized integrating advanced features that simplify the overall operation. Through these enhancements, the product will better meet user needs while maintaining its core purpose. The ultimate goal was to create a product that is not only functional but also intuitive and accessible to a broader audience.

The design stage focused on determining and selecting the appropriate formulas that aligned with the objectives to be achieved in the process. At this stage, careful planning and formulation were carried out to design specific formulas that will be implemented in the item analysis application. These formulas played a crucial role in facilitating users to analyze and interpret data related to the quality of test items effectively. By utilizing these formulas, users can gain insights into whether the questions are valid and reliable for use in assessments or examinations. The design ensured that the application provides accurate and comprehensive results, making the evaluation process more efficient. Additionally, the formulas were tailored to address the specific needs of the users, ensuring they align with best practices in educational measurement. Thus, this stage layed a solid foundation for creating a tool that supports the improvement of test item quality in various educational settings.

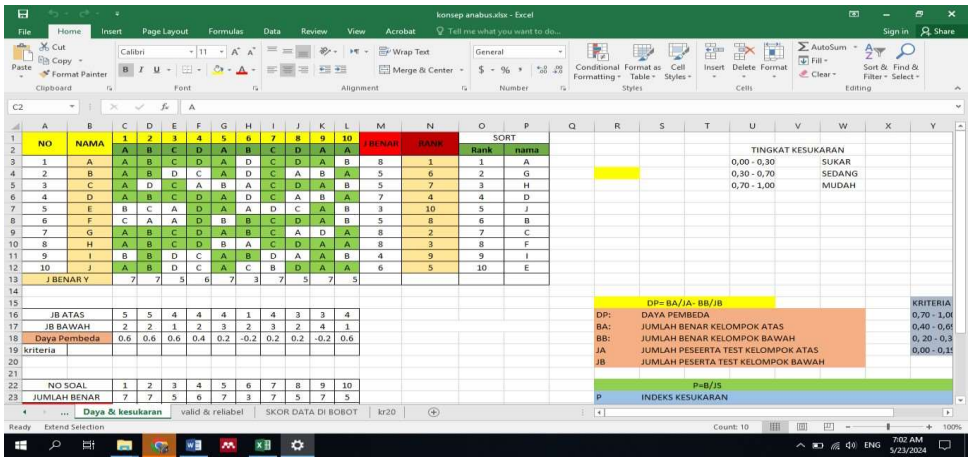

Figure 1. Formula Design

The figure above illustrated the preliminary design of a formula created to calculate the results of the multiple-choice item analysis process. This formula was initially developed and implemented in Excel to facilitate data computation and analysis. At this stage, the results obtained from the formula are carefully reviewed to assess their accuracy and relevance. The evaluation process aimed to identify any potential areas for refinement or optimization to enhance the reliability of the analysis. Below, the display of the calculation results was presented for further observation and improvement.

Figure 2. Display of Calculation Results

The calculation results menu, which was originally designed in Excel, has undergone a significant transformation to better meet user needs. This redesign introduced a more user-friendly interface that simplifies the process of reviewing and understanding calculation results. By enhancing the layout and functionality, users can more easily navigate and interpret the data they input into the Anabus application. The improvements aimed to provide a seamless experience, reducing complexity and making the application accessible to a broader audience. Ultimately, this transformation enhanced usability and ensures that users can efficiently analyze the results of their calculations.

Development and Implementation Stage

The development stage involved a series of systematic activities designed to ensure the quality and functionality of the resulting product. These activities included collecting a wide range of relevant sources to enrich the formulation of methods for determining the quality of items, ensuring a comprehensive and reliable basis for evaluation. Additionally, the development process emphasized the creation of necessary displays and generating export features that facilitate users in accessing and interpreting the results effectively. This stage also included validating the initial draft of the development products through expert reviews, which serve as a critical step in identifying and addressing potential weaknesses. The feedback obtained from these expert evaluations was then used as a basis for revisions, ensuring that the final product meets the required standards of quality and usability.

The development of a multiple-choice item analysis application has reached the stage of feasibility assessment, involving evaluations by media experts and teachers through detailed questionnaires. The primary aim of the assessment by media experts was to provide valuable feedback for refining and improving the application, ensuring it meets professional standards and user expectations. On the other hand, response questionnaires distributed to individual teachers serve to collect comprehensive input and suggestions regarding the effectiveness, usability, and overall functionality of the Anabus application, designed for multiple-choice item analysis. These evaluations were essential for identifying strengths and areas for improvement in the application's interface and performance. Furthermore, the application has been designed to

display detailed calculation results, offering clear and user-friendly insights into the analysis process. This feature ensured that users can easily interpret data and make informed decisions based on the application's output. By combining expert reviews with user feedback, the development process aimed to create a robust and practical tool tailored to the needs of educators.

Following the development of the Anabus application (multiple-choice item analysis), the next step was to validate the produced products. The following described the product trial activities for the question item analysis application, starting with trials by media experts. The results of these trials, as assessed by media experts, were summarized in Table 1. The percentage of validation by the three media experts can be calculated as follows.

Table 1. Media Expert Validation Results

Media Expert Validation	Feasibility Percentage	Criteria
First media expert	85%	Very feasible
Second media expert	86, 66%	Very feasible
Third media expert	95%	Very feasible

In the development of the Anabus application, which was designed for the analysis of multiple-choice test items, the subsequent stage involves validating the products that have been developed. This validation process was essential to ensure the reliability and effectiveness of the application. The validation began with trials conducted by media experts, whose expertise is crucial in evaluating the usability, design, and functionality of the application. These media experted thoroughly assess various aspects of the application, providing detailed feedback to refine its features.

Table 2. Teacher Response Results

Teacher Response Percentage	Feasibility Percentage	Criteria
Teacher Response	80,5%	Worth

The results of their evaluations were then compiled and summarized in Table 2, which provided a comprehensive overview of their findings. To determine the overall effectiveness of the application, the percentage of validation achieved by the three media experts was calculated systematically. This calculation serves as a quantitative measure of the application's readiness for further development and broader implementation.

Evaluation Stage

The final stage involved conducting an evaluation, which included analyzing and making final revisions to address the deficiencies identified in the Anabus application (multiple-choice item analysis) developed based on suggestions from the assessment questionnaire. The goal was to ensure that the developed item analysis application (Anabus) is suitable for use across a broader range of schools and universities. The product revision stages based on criticism and suggestions from media experts included the following. In this main menu,

the cover design was not in accordance with the function of the application, so it is not interesting if it was placed on the cover display in the Anabus application. The following was a picture of the product display before revision and after revision.

Figure 3. Cover menu before revision

The following menu showcased a revision of the initial main menu, which was previously deemed unattractive and lacked aesthetic appeal. This redesign has been carefully crafted based on feedback and constructive critiques provided by media experts. The aim of this improvement was to create a visually appealing and user-friendly interface that aligns seamlessly with the overall design of the application. A well-designed and engaging interface was a crucial factor in capturing the attention of users and fostering their interest in exploring the features of the Anabus application. By incorporating expert suggestions, the updated menu has been enhanced to provide a more consistent and cohesive user experience. The revised design prioritized both functionality and aesthetics, ensuring that it meets user expectations while maintaining alignment with the application's branding. The following section presents a detailed illustration of the updated menu's appearance, highlighting its key features and improvements.

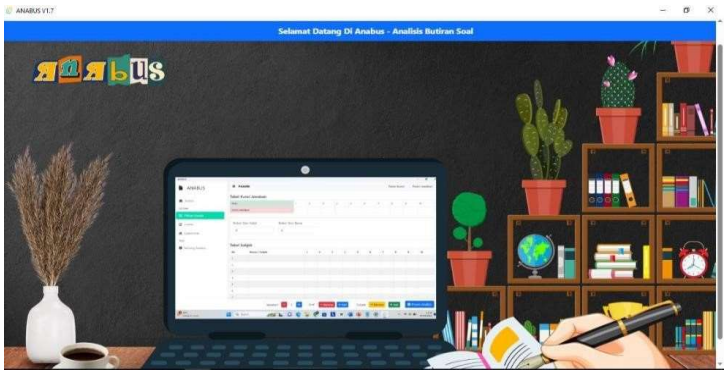

Figure 4. Cover Menu After Revision

In this second iteration of the menu, the presentation of information about the Anabus application lacks structure and clarity, making it challenging for users to grasp its key features and functionalities. The absence of well-organized and detailed explanations results in confusion and leaves users unable to understand the purpose or benefits of the application. To address this issue, it was crucial to provide a thorough and comprehensive description of the Anabus application, including a clear outline of its objectives, primary functions, and

user advantages. Such details were essential to ensure that users can fully comprehend what the application offers and how it can meet their needs. A well-structured presentation should include a concise purpose statement, detailed feature descriptions, and examples of practical use cases. Below was an illustration of the product's appearance before incorporating the necessary revisions for improvement.

Figure 5. Information Menu About Anabus Before Revision

The following menu was a revised version of the second menu, which initially lacked complete information about the Anabus application. In this revision, the menu has been improved by adding comprehensive details about Anabus, addressing the gaps present in the previous version. This revision was made based on feedback, constructive criticism, and suggestions from media experts to ensure the information provided is more accurate and user-friendly. The enhancements aimed to make the description of Anabus clearer and more accessible for application users. By incorporating these revisions, users can now easily understand the purpose, features, and functionality of the Anabus application. The updated product display reflected these improvements, offering a more detailed and engaging experience for users. This revision demonstrated a commitment to addressing user needs and providing high-quality information. Here was the updated product display after the revision.

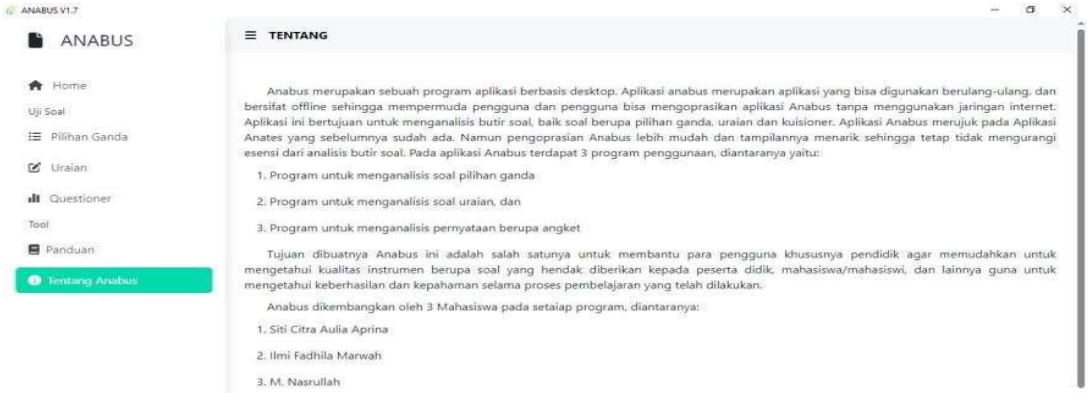

Figure 6. Information Menu About Anabus After Revision

On this third menu is the Anabus application guide. On this menu, the application usage guide has not been included in the application, so users have difficulty if they want to operate the application. The application usage guide should be included in the guide menu contained in the application, to make it easier for users to operate the Anabus application. The following is a picture of the product display before revision.

Figure 7. Anabus Guide Menu Before Revision

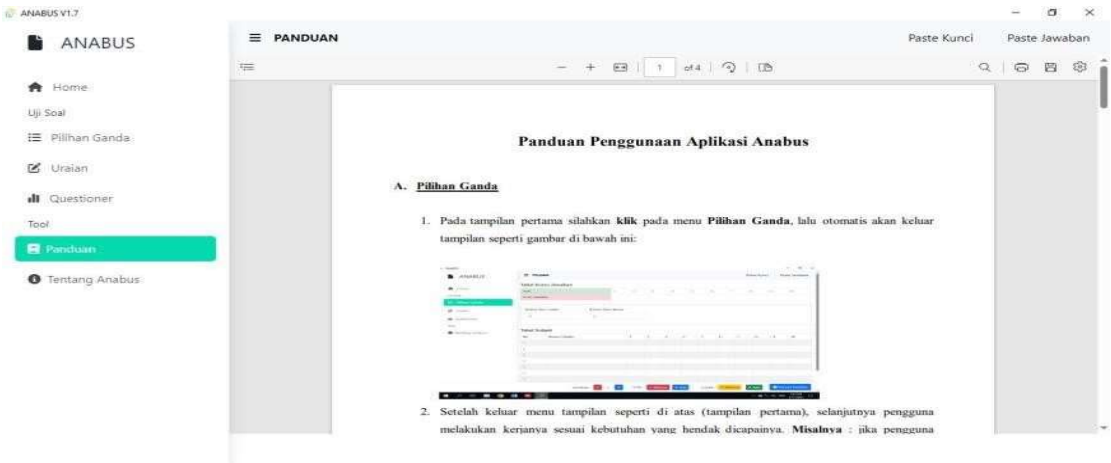

Figure 8. Anabus Guide Menu Before Revision

The menu display below represented a revised guide for the Anabus application. Initially lacking a usage guide, feedback from media experts during validation highlighted the necessity of including a clear application guide. This enhancement aimed to improve user accessibility and satisfaction with the Anabus application. The following section presented the updated menu display.

Discussion

The research results indicated that the *Anabus* application (Multiple-Choice Item Analysis) developed using the ADDIE model successfully met the eligibility criteria as a tool for analyzing multiple-choice items. In educational assessment, analyzing multiple-choice items was critical for evaluating the effectiveness and fairness of test questions. Traditional methods often involved manual

calculations or the use of basic software like spreadsheets, which might be limited in functionality, user accessibility, and scalability. Anabus's adherence to the ADDIE model aligned with a well-established instructional design framework consisting of Analysis, Design, Development, Implementation, and Evaluation. This model ensured a systematic, data-driven, and iterative approach to educational tool development, guaranteeing alignment with learning objectives and continuous improvement of the product's efficacy and user experience. The ADDIE model was widely recognized for facilitating structured course design, customization, collaboration among stakeholders, and continuous enhancement (Abuhassna et al., 2024).

Advanced statistical tools such as Item Response Theory (IRT) or Classical Test Theory (CTT) were widely recognized but often require specialized knowledge and resources, making them less accessible to general educators. Item analysis was grounded in classical psychometric theories such as Classical Test Theory (CTT) and Item Response Theory (IRT). CTT focused on test validity and reliability by evaluating observed scores as composites of true scores and measurement errors, thus emphasizing consistency and accuracy in assessment tools. Key indices used for item evaluation include difficulty level, discrimination power, and distractor effectiveness, which align with the functionalities offered by Anabus for multiple-choice analyses (Beerepoot, 2023; Sicilia et al., 2022). IRT provided a more sophisticated probabilistic framework to model the relationship between an individual's latent ability and item characteristics. It allowed detailed item-level evaluation and supports applications like adaptive testing and detecting item bias, facilitating high-precision assessment development. Integrating these statistical methods into Anabus enhanced its technical rigor while maintaining usability (Crawford et al., 2023; Lahza et al., 2023).

This research addressed a gap in existing tools by developing a dedicated desktop application, Anabus, which simplified item analysis while ensuring accuracy and user-friendliness. The application leveraged modern design principles to integrate complex statistical functions into an intuitive interface, making it accessible for educators with varying levels of technical expertise.

The *Anabus* application standed out due to several key innovations: 1) Integration of ADDIE Model, while many educational tools are developed without structured frameworks, *Anabus* followed the ADDIE model, ensuring systematic development through stages of Analysis, Design, Development, Implementation, and Evaluation. This methodology enhanced the reliability and adaptability of the final product. 2) User-Centric Design, the application's interface was tailored to educators, emphasizing simplicity, clarity, and efficiency. The transformation of initial designs from Excel formulas to a fully functional desktop application illustrates its usability focus (Kadaskar, 2024; Y. Liu et al., 2024). 3) Comprehensive Validation, validation by media experts and educators demonstrated a robust approach to ensuring quality. Achieving feasibility ratings above 80% confirmed the application's practical value and readiness for implementation. 4) Bridging the Gap, by combining advanced item

analysis capabilities with user-friendly features, *Anabus* bridges the gap between complex statistical tools and the operational needed of educators in real-world settings.

Compared to existing tools, *Anabus* offered a unique balance of technical sophistication and accessibility. Tools such as SPSS, R, or dedicated IRT software focus heavily on statistical rigor but often lack user-friendly interfaces, requiring advanced knowledge to operate effectively (Garcia et al., 2022). On the other hand, simpler tools might sacrifice accuracy or functionality for ease of use. Anabus innovates by 1) Providing built-in formulas for item validity and reliability, tailored for multiple-choice analysis. 2) Offering exportable results and customizable interfaces to suit diverse educational contexts. 3) Including detailed user guided and information menus, ensuring usability without prior expertise in statistical methods (Harahap & Ritonga, 2023; Nurbayan & Anwar, 2022).

The Anabus application significantly impacted learning by enabling teachers to assess the quality of their multiple-choice test items effectively. study by Karim et., al, item analysis revealed that many test questions had poor or mediocre quality in terms of difficulty and discriminating power, which often goes unnoticed without proper analysis. By using Anabus, teachers can identify low-quality questions that need revision, ensuring more accurate assessments and improving students' learning outcomes. This tool empowerd educators to enhance test quality, ultimately supporting better teaching and learning practices (Karim et al., 2021). The introduction of *Anabus* had significant implications for educational institutions: 1) Improving Assessment Quality: By automating and simplifying item analysis, *Anabus* enabled educators to identify and refine low-quality questions, ensuring that assessments align with learning objectives and standards. 2) Professional Development: The tool serves as a learning platform for educators to better understand assessment principles, contributing to their professional growth. 3) Broad Accessibility: Designed for desktop use, *Anabus* was a cost-effective solution for schools and universities with limited access to advanced statistical software or technical training.

The *Anabus* application represented a significant advancement in the field of educational assessment tools. By addressing the limitations of existing software and prioritizing user needs, it enhanced the accessibility and quality of item analysis for educators. As an innovative, scalable, and user-friendly tool, *Anabus* had the potential to transform assessment practices and contribute to the broader goal of improving educational outcomes globally. Future developments aimed at expanding its features and adaptability will further solidify its role as *a state-of-the-art* solution in educational technology.

CONCLUSION

The results of the study prove that the Anabus application, developed using the ADDIE model, is feasible and effective as a tool for analyzing multiple-choice questions in educational assessments. The study shows that a systematic, user-centered approach to application development can address teachers' need

for accurate, user-friendly analysis tools. This study's main contribution is providing a technological solution that improves the quality of item analysis and supports teachers' professional development in understanding learning evaluation principles. However, the scope of features is limited, and more testing is needed at various levels of education to make the application more adaptive. Further research is recommended to develop more flexible application features and expand trials to increase Anabus's usefulness in various educational contexts.

ACKNOWLEDGMENT

The author would like to express his deepest gratitude to the Syekh Wasil Kediri Postgraduate Program and KH. A. Wahab Hasbullah University Jombang for their support and facilities provided during the writing of this article. The assistance, guidance, and opportunities provided greatly contributed to the smooth running of the research and writing of this article. May this collaboration continue and bring benefits to the development of science and education.

REFERENCES

Abuhassna, H., Alnawajha, S., Awae, F., Adnan, M., & Edwards, B. I. (2024). Synthesizing Technology Integration Within the Addie Model for Instructional Design: a Comprehensive Systematic Literature Review. *Journal of Autonomous Intelligence*, *7*(5), 1–28. 10.32629/jai.v7i5.1546

Astiza, D. A., Hidayatulloh, M. K. Y., & Ashoumi, H. (2023). The Influence of the SQ4R Learning Model on Learning Outcomes Student. *APPLICATION: Applied Science in Learning Research*, *3*(2), 33–37. https://doi.org/10.32764/application.v3i2.4750

Balasopoulou, A., Kokkinos, P., Pagoulatos, D., Plotas, P., Makri, O. E., Georgakopoulos, C. D., Vantarakis, A., Li, Y., Liu, J. J., Qi, P., Rapoport, Y., Wayman, L. L., Chomsky, A. S., Joshi, R. S., Press, D., Rung, L., Ademola-popoola, D., Africa, S., Article, O., … Loukovaara, S. (2017). Symposium Recent Advances and Challenges in the Management of Retinoblastoma Globe-Saving Treatments. *BMC Ophthalmology*, *17*(1), 1. https://doi.org/10.4103/ijo.IJO

Beerepoot, M. T. P. (2023). Formative and Summative Automated Assessment with Multiple-Choice Question Banks. *Journal of Chemical Education*, *100*(8), 2947–2955. https://doi.org/10.1021/acs.jchemed.3c00120

Belay, L. M., Sendekie, T. Y., & Eyowas, F. A. (2022). Quality of Multiple-Choice Questions in Medical Internship Qualification Examination Determined by Item Response Theory at Debre Tabor University, Ethiopia. *BMC Medical Education*, *22*(1), 1–11. https://doi.org/10.1186/s12909-022-03687-y

Belkhir, F. Z. (2024). Challenges and Opportunities of AI-Assisted Learning: A Systematic Literature Review on the Impact of ChatGPT Usage in Higher Education. *IJLTER*, *22*(7), 25–39. https://doi.org/https://www.nature.com/articles/s41586-023-06221-2

Chin, H., Chew, C. M., Lim, H. L., & Thien, L. M. (2022). Development and Validation of a Cognitive Diagnostic Assessment with Ordered Multiple-Choice Items for Addition of Time. *International Journal of Science and Mathematics Education*, *20*(4), 817–837. https://doi.org/10.1007/s10763-021-10170-5

Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is Needed for Ethical Chatgpt: Character, Assessment, and Learning Using Artificial Intelligence (AI). *Journal of University Teaching and Learning Practice*, *20*(3), 1–21. https://doi.org/10.53761/1.20.3.02

Darvin, R., & Norton, B. (2023). Investment and Mmotivation in Language Learning: What's the Difference? *Language Teaching*, *56*(1), 29–40. https://doi.org/10.1017/S0261444821000057

Forsblom, L., Pekrun, R., Loderer, K., & Peixoto, F. (2022). Cognitive Appraisals, Achievement Emotions, and Students' Math Achievement: A Longitudinal Analysis. *Journal of Educational Psychology*, *114*(2), 346–367. https://doi.org/10.1037/edu0000671

Garcia, A. R., Filipe, S. B., Fernandes, C., Estevão, C., & Ramos, G. (2022). Detection of Bipolar Disorder in the Prodromal Phase: a Systematic Review Of Assessment Instruments. *Study Selection Criteria*, *1*(1), 1–79. https://doi.org/https://www.nature.com/articles/s41586-023-06221-2

Halverson, L., Graham, C. R., & Henrie, C. (2020). Learning, Design, and Technology. in *Learning, Design, and Technology* (Issue February 2021). https://doi.org/10.1007/978-3-319-17727-4

Harahap, H. N., & Ritonga, A. A. (2023). The Effect of Using Project Based Learning Model with Autoplay Media Studio on Learning Achievement. *Edureligia: Jurnal Pendidikan Agama Islam*, *07*(01), 99. https://doi.org/10.1016/j.mjafi.2020.11.007

Hidayatulloh, M. K. Y., & Ashoumi, H. (2022). Creativity and Entrepreneur Knowledge to Increase Entrepreneurial Intent Among Vocational School Students. *Journal of Education and Learning (EduLearn)*, *16*(4), 434–439. https://doi.org/10.11591/edulearn.v16i4.19771

Ismail, M. I. (2020). *Asesmen dan Evaluasi Pembelajaran* (1st ed.). Cendekia Publisher.

Kadaskar, H. R. (2024). Enhancing User Experience in Mobile Application Design Through Gestural Interaction: a Human-Computer Interaction Perspective. *International Journal of Scientific Research in Modern Science and Technology*, *3*(8), 1–6. https://doi.org/10.59828/ijsrmst.v3i8.239

Karim, S. A., Sudiro, S., & Sakinah, S. (2021). Utilizing Test Items Analysis to Examine the Level of Difficulty and Discriminating Power in a Teacher-Made Test. *EduLite: Journal of English Education, Literature and Culture*, *6*(2), 256. https://doi.org/10.30659/e.6.2.256-269

Khoirunnisa, S. (2020). *Aplikasi Penilaian Kinerja Soal Pilihan Ganda Dengan Microsoft Excel 2013*.

Kumar, D., Jaipurkar, R., Shekhar, A., Sikri, G., & Srinivas, V. (2021). Item Analysis of Multiple Choice Questions: a Quality Assurance Test for an Assessment Tool. *Medical Journal Armed Forces India*, *77*(1), S85–S89. https://doi.org/10.1016/j.mjafi.2020.11.007

Laato, S., Morschheuser, B., Hamari, J., & Bjorne, J. (2023). AI-Assisted Learning with ChatGPT and Large Language Models: Implications for Higher Education. *Proceedings - 2023 IEEE International Conference on Advanced Learning Technologies, ICALT 2023*, *12*(November 2022), 226–230. https://doi.org/10.1109/ICALT58122.2023.00072

Lahza, H., Smith, T. G., & Khosravi, H. (2023). Beyond Item Analysis: Connecting Student Behaviour and Performance Using E-Assessment Logs. *British Journal of Educational Technology*, *54*(1), 335–354. https://doi.org/10.1111/bjet.13270

Liu, W. (2021). Does Teacher Immediacy Affect Students? A Systematic Review of the Association Between Teacher Verbal and Non-verbal Immediacy and Student Motivation. *Frontiers in Psychology*, *12*(June), 1–13. https://doi.org/10.3389/fpsyg.2021.713978

Liu, Y., Tan, H., Cao, G., & Xu, Y. (2024). Enhancing User Engagement Through Adaptive UI/UX Design: a Study on Personalized Mobile App Interfaces. *Computer Science & IT Research Journal*, *5*(8), 1942–1962. : 10.53469/wjimt.2024.07(05).01

Mai, D. T. T., Da, C. Van, & Hanh, N. Van. (2024). The Use of ChatGPT in Teaching and Learning: a Systematic Review Through SWOT Analysis Approach. *Frontiers in Education*, *9*(February), 1–17. https://doi.org/10.3389/feduc.2024.1328769

Mesra, R. (2023). Research & Development dalam Pendidikan. In *PT. Mifandi Mandiri Digital*. PT. Mifandi Mandiri Digital.

Nurbayan, Y., & Anwar, S. (2022). Digital Library Utilization ; Strategies to Improve Digital Islamic Digital Literacy for Religion Teachers. *Edureligia: Jurnal Pendidikan Agama Islam*, *6*(December), 150–160. https://doi.org/10.33650/edureligia.v6i2.4536

Peters, S. J. (2022). The Challenges of Achieving Equity Within Public School Gifted and Talented Programs [University of Wisconsin – Whitewater]. In *Gifted Child Quarterly* (Vol. 66, Issue 2). https://doi.org/10.1177/00169862211002535

Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Al Ameer, A. Y., Yahia, A. I. O., Mohammed, O. A., & Adam, M. I. E. (2024). Item Analysis: the Impact of Distractor Efficiency on the Difficulty Index and Discrimination Power of Multiple-Choice Items. *BMC Medical Education*, *24*(1), 1–7. https://doi.org/10.1186/s12909-024-05433-y

Ribeiro-Silva, E., Amorim, C., Aparicio-Herguedas, J. L., & Batista, P. (2022). Trends of Active Learning in Higher Education and Students' Well-Being: A Literature Review. *Frontiers in Psychology*, *13*(April), 1–10.

https://doi.org/10.3389/fpsyg.2022.844236

Rozi, F., Zahiro, R., & Hotimah, A. S. (2022). Training on the Utilization of Used Goods as APE Based on Natural Potential. *Indonesian Journal of Community Research & Engagement*, *01*(01), 1–7. https://doi.org/http://doi.org/10.15408/ijocore.vxix.xxxx

Sicilia, A., Alcaraz-Ibáñez, M., Paterna, A., & Griffiths, M. D. (2022). A Review of the Components of Problematic Exercise in Psychometric Assessment Instruments. *Frontiers in Public Health*, *10*(March), 1–13. https://doi.org/10.3389/fpubh.2022.839902

Slepkov, A. D., Van Bussel, M. L., Fitze, K. M., & Burr, W. S. (2021). A Baseline for Multiple-Choice Testing in the University Classroom. *SAGE Open*, *11*(2), 1–12. https://doi.org/10.1177/21582440211016838

Vaughn, S. R., Bos, C. S., & Schumm, J. S. (2023). *Teaching Students Who are Exceptional, Diverse, and at Risk in the General Educational Classroom* (pp. 1–7). Pearson Education. https://doi.org/10.1016/j.tate.2022.103945

You, K., Liu, Y., Wang, J., & Long, M. (2021). LogME: Practical Assessment of Pre-trained Models for Transfer Learning. *Proceedings of Machine Learning Research*, *139*(1), 12133–12143. https://proceedings.mlr.press/v139/you21b.html