Research on the Development of Assessment Instruments in Indonesia

A Critical Literature Review

M. Rizky Satria

Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
rizkysatria.2022@student.uny.ac.id

Abstract. Education development in the second decade of the 21st century leads to attaining 21st-century skills. Indonesia responded to its outgrowth through a series of curriculum reforms to achieve student competencies to face world challenges. This study aims to provide an overview of how the research on developing assessment instruments in Indonesia has followed these developments over the past few years. A total of 442 articles were collected with the keywords of *instrument development, assessment,* and *evaluation* from the Scopus & Google Scholar database from 2021 to mid-2023. The literature review analyzes data based on a thematic model that reveals school levels, related subjects, and assessment targets. The findings show that research on assessment development in Indonesia was largely dominated by the high-school level, the sciences group, and the cognitive target. This review then discusses critically how research on the development of assessment needs to move beyond the partial aspect towards a more comprehensive and integrative approach to optimize the achievement of expected student competency.

Keywords: Assessment, Competency, Instrument Development.

1 Introduction

Assessment is an important component of education. A good assessment will encourage teachers to determine appropriate teaching strategies. Therefore, improving the quality of education requires improvements in the assessment system [1]. In Indonesia, assessment standards refer to the Regulation of the MoECRT No. 21 of 2022 concerning Educational Assessment Standards which explains that assessment is the process of collecting and processing information to determine learning needs and developments or learning outcomes of students [2].

Entering the 21st century, the assessment system faces challenges due to the increasing educational needs including achieving the latest competencies. These competencies include the skills that are needed to enter the digital era and the development of information technology [3]. Over the country, this development was responded by a series of national curriculum reforms that shifted the focus from content mastery to competency mastery, starting with the emergence of the Competency-Based Curriculum in 2004. In addition to focusing on achieving student competencies, the education sector in Indonesia is also starting to seriously develop a

P. C. Kuswandi et al. (eds.), *Proceedings of the 6th International Conference on Current Issues in Education (ICCIE) 2023*, Advances in Social Science, Education and Humanities Research 847,
https://doi.org/10.2991/978-2-38476-245-3_9

focus on character achievement with the existence of Regulation of President No. 87 of 2017 concerning Strengthening Character Education, which was followed up by the Regulation of MoECRT No. 20 of 2018 concerning Strengthening Character Education in Formal Education. The Character Education Program is implemented by applying values such as being religious, honest, tolerant, disciplined, working hard, creative, independent, democratic, having curiosity, nationality, patriotism, respecting achievements, communicative, love of peace, love of reading, environmental care, social care, and responsibility [4].

The momentum for optimizing education which is centered on achieving student character and competency occurred with the National Assessment program in 2021 which surveyed student character and competency. The results were released one year later and became a database for designing learning activities with identified targets of character and competency that still needed to be improved. However, to achieve the expected character and competency targets, appropriate assessment tools are needed. So that research regarding the development of up-to-date assessment instruments is a necessity. This is important because the field findings show that teachers and schools still have difficulty to develop and, especially, to evaluate how students achieve the expected competencies and character. One of the main factors is that teachers do not have sufficient references regarding developing assessment instruments beyond the students' cognitive abilities [5], [6].

Referring to the importance of providing reference assessment instruments for teachers, there is no study yet to look at the overall development of research related to the development of assessment instruments in Indonesia. Therefore, this study aims to provide an overview of research trends related to the development of assessment over the past few years and to present a critical review of how the research is relevant to the latest education development.

2 Method

This study uses a literature review method which was elaborated critically to examine research results regarding the development of student assessment instruments in Indonesia from 2021 to mid-2023 (January 2021-August 2023).

The first step is to collect data obtained from the Scopus and Google Scholar databases using the Publish or Perish application to find all academic reports, including journal articles, proceedings, and student theses or dissertations with the keywords: instrument development, assessment, and evaluation. The type of research includes needs analysis, instrument analysis, and instrument development with various types of methods.

The second step is to conduct selection. In the initial search, 549 articles were collected, dominated by journal articles, both national and international journals. Data sorting was then carried out by eliminating articles with the following criteria: (1) There was duplication, whether two of the same articles appeared or one study appeared in several articles, (2) The research context is not for students at school (Excluding research on university students, teachers, school principals, or other roles outside the field of education), (3) Written work in the form of books, modules, or other types that are not the research, (4) Content is about training activities or

socialization on instrument development, (5) Articles whose sources cannot be traced or have the "retracted" status. After sorting was complete, 442 research works were collected which were ready to be processed.

The third step is curation. Using a matrix processed with Microsoft Excel software, all data is curated based on the thematic model that reveals school levels, related subjects, and assessment targets. The fourth step is to analyze all the data to find patterns or findings related to the research objectives. Meanwhile, the fifth step is to conduct a critical review of the findings obtained in the previous stage. The process in the last three steps was conducted interactively.

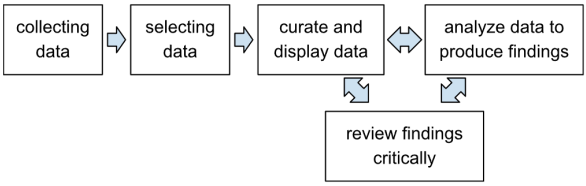

Fig. 1. Data processing flow

3 Result

The findings are grouped into 3 main categories: school-level context of research, related subject, and assessment target.

3.1 School-Level Context of Research

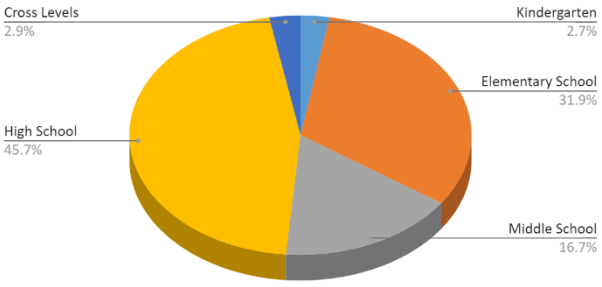

Fig. 2. School-level context

The first categorization produces a mapping of research context: Research on instrument development is mostly conducted at the high school level (45.7%), followed by elementary school (31.9%), middle school (16.7%), cross levels (2.9%), and kindergarten (2.7%). There are two criteria for the cross-level, related to research

conducted at all levels (e.g., the development of assessment instruments for self-confidence) and research conducted at special education schools (e.g., instruments for assessing functional residual vision or writing ability for students with cerebral palsy). The initial finding of why the high school level has the largest number of studies is because high school teachers have a more varied composition with the specialty of one teacher for one subject so the college majors for high school teachers are more numerous and varied. Therefore, this mapping data shows reasonable conditions.

3.2 Related Subject

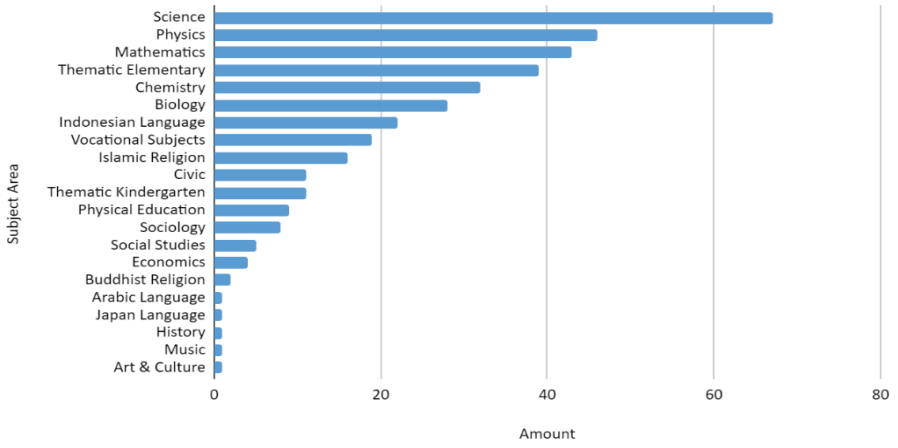

Fig. 3. Related subject

The second categorization produces subjects related to the development of assessment instruments at school. In the first place, Science refers to integrated science studied in elementary and middle school, while subjects such as Physics, Chemistry, and Biology are studied separately in high school. The data shows that research on assessment is dominated by a group of natural sciences rather than the social sciences. The general estimate regarding this condition is that research on instrument development which is related to quantitative processing is more interesting to natural science researchers or educators than to social science researchers or educators. However, students as assessment subjects should recei ve the same quality of assessment in all the subjects they study. Therefore, this condition needs to be a concern for the stakeholders to encourage more equitable research on all subjects at school.

3.3 Assessment Target

Table 1. Assessment target

Assessment Target	Amount	Percentage
Knowledge	151	34.0
Skills	115	26.0
Attitude	71	16.1
Literacy	38	8.6
Character	22	5.0
Critical Thinking	35	7.9
Creativity	8	1.8
Collaboration	15	3.4
Communication	11	2.5
Problem Solving	9	2.0
Inquiry	1	0.2

The third categorization produces data mapping assessment targets on aspects of knowledge, skills, and attitudes that are partially developed which occupy the third highest compared to other aspects related to literacy, character, critical thinking, creativity, collaboration, communication, problem-solving, and inquiry skills. This mapping is extensive, allowing one study to have several measurement targets. For example, research by Syanas entitled "Development of a Computer-Based Testlet Assessment Instrument to Measure Critical Thinking Skills and Scientific Literacy" has targeted critical thinking and literacy, while research by Shintya entitled "Development of an Assessment Instrument to Measure Collaboration and Problem-Solving Skills in Physics Learning" has a target of collaboration and problem-solving. Although this data mapping is not rigid, it can be seen how much each assessment target appears in percentage compared to the total number of studies being analyzed.

Regarding the three major assessment targets identified, the Knowledge category is dominated by High Order Thinking Skills/HOTS measurements (83), followed by material understanding (56), and creative thinking (12) which is differentiated from creativity as a more complete competency. The Skills category is a combination of psychomotor domain measurements (19) and various types of performance (96). Psychomotor measurements are dominated by kindergarten level, physical education subjects, and science laboratory material. Meanwhile, performance measurement is dominated by vocational levels and science groups. The Attitude category includes various forms that are closely related to character, such as social attitudes, responsibility, discipline, scientific attitudes, morality, environmental concern, etc. but only develop in the limited scope of the affective domain.

Other assessment targets outside the focus on knowledge, skills, and attitudes only constitute less than 10% of the total research results collected.

4 Discussion

4.1 The Partial Aspect of Assessment

There are two findings related to partial concern, that research on assessment has not proportional to all subjects, especially the social science group, and to all aspects of ability, especially aspects of character and competency which require a combination of knowledge, skills, and attitudes. Considering that each subject has the same importance, assessment research should receive the attention of stakeholders in all learning to accommodate the instrument characteristics of each subject. Assessment research in social sciences subjects needs to be increase and expand so that the assessment process can take place optimally in all subjects taken by students because basically, every teacher needs to optimize assessments in their classrooms to be able to optimize the learning process [7].

As important as ensuring assessment research is equitable across all subjects, teachers and researchers also need to ensure assessment research is equitable across all aspects of student abilities. Data result that assessment research is still dominated by the knowledge aspect shows that the research trends have not developed proportionally. The challenge in this regard is the view that students' important abilities are not only in the knowledge aspect, but also in attitudes and skills, including character and competence which combine all of them [8].

Another finding from the data collected was the tendency of researchers to measure aspects of Knowledge, Skills, and Attitudes (KSA) separately, considering that these three aspects are the most popular assessment elements in education. From a regulatory perspective in Indonesia, the Graduate Competency Standards published in 2016 explains that graduate school competency is divided into these three aspects. From the methodological perspective, the assessment of student ability is divided into different measurement domains generally known as cognitive, affective, and psychomotor [9] or cognitive and non-cognitive [1]. These domains, when formulated into learning objectives, turn into knowledge, skills, and attitudes.

According to the separation of KSA, there are three issues that need attention, including: First, the KSA aspects are still being studied partially, rather than forming an integrated whole. Second, the KSA is viewed in an unfair manner with a greater focus on knowledge or cognitive aspects. Third, the pattern of separating these three aspects, which focuses on the knowledge aspect, making it difficult for educators to measure student competence as a whole. This condition moves away from the needs of the times which see that competency is a unity of KSA aspects to do a particular job [10]. Even the character does not contain one aspect of attitude or affective domain, but rather a collection of elements of KSA (moral knowing, moral feeling, and moral action) which are integrated and form a unity [11].

4.2 Towards the Comprehensive and Integrative Assessment

The current trend in assessment development is towards 21st-century skills that are not limited to a few subjects and are a combination of comprehensive aspects of student abilities, such as assessments for literacy, numeracy, problem-solving,

collaboration, communication, etc. [12], [13]. Therefore, more assessment research is needed that is more comprehensive, evenly distributed across all subjects, even coordinative across subjects, and goes beyond assessment targets that only achieve knowledge aspects or cognitive domains. Measuring academic success is not always relevant only by measuring cognitive abilities that produce grades, so a more comprehensive instrument targeting a combination of knowledge, attitudes, and skills is necessary [14].

Traditional assessments that rely on cognitive abilities are no longer optimal in measuring student competencies in the current era. [12] Therefore, it is necessary to develop instruments that are more appropriate to current developments. However, changing assessment targets beyond the aspects of KSA is often challenging because educators have to build new forms of assessment [13]. Thus, change must take place systemically from the policy level to implementation in the field. In Indonesia, the government has made systemic changes starting from a regulatory perspective by no longer using the "trichotomy" of knowledge, skills, and attitudes in updating the Regulation of Graduate Competency Standards and Educational Assessment Standards in 2022. Then in terms of implementation, through a curriculum guide document, a mechanism for implementing learning and assessment has been outlined with greater emphasis on the process by providing more space for teachers to explore various forms of comprehensive assessment [15]. This context should become the cornerstone for the development of more comprehensive assessment instruments that transcend partial aspects of ability.

Move beyond regulatory grounds, researchers, teachers, and educators can re-elaborate on methodologies that have been designed by experts over the last two decades, such as the 21st-century skills assessment methodology developed by Marzano & Heflebower [16] based on a combination of cognitive and conative skills, Greenstein [17] which presents a framework for measuring thinking skills, actions, and Living Skills through multipurpose assessments, or Wiggins [18] who promotes the comprehensive approach through authentic assessment. However, the regulation and methodological basis can be a strong rationale for subsequent research to develop the expected assessment instrument.

5 Conclusion

This literature study has provided an overview of research trends in the development of assessment instruments in Indonesia over the past few years, which have been dominated by the high-school level, the sciences group, and the cognitive target. The concern then arises at the second and third points, where there is inequality in the distribution of research on lesson subjects in schools and inequality in assessment targets or measured abilities. These disparities can be overcome by moving towards a more comprehensive and integrative assessment trend. Comprehensive includes an even distribution of all lessons and assessment targets, while integrative combines KSA to measure character and competency.

Ultimately, this review provides a background for researchers to develop assessments across subjects proportionally and move beyond cognitive abilities toward more integrative competencies in line with current development needs. On the

other hand, this study also shows that assessment development targeting character and competence may become challenging, but according to the updated regulations and scientific methodology, the shift in instrument research trends from partiality to comprehensive and integrative assessment has received the rationalization to be developed.

6 Limitation

This study still has data coverage limitations. Not all research results, especially student theses and dissertations, can be collected because they are not included in the Scopus and Google Scholar databases or have not been published yet. This research also has a limited time span and context coverage. Hopefully, further research can be conducted to gather a wider range of data and research context to advance research in the field of assessment instrument development.

Authors' Contributions. The author confirms sole responsibility for the study conception and design, data collection, and analysis of the results.
**Research data: https://bit.ly/4aMeRqo*

Acknowledgments. The authors would like to thank the Master Program in Educational Research and Evaluation, Universitas Negeri Yogyakarta for providing facilities, access, and support so that this article could be produced.

References

1. D. Mardapi, Pengukuran, Penilaian, dan Evaluasi Pendidikan, Parama Publishing, 2019.
2. Regulation of the Minister of Education, Culture, Research, and Technology of the Republic of Indonesia Number 21 of 2022 concerning Educational Assessment Standards. [Online]. Available: https://jdih.kemdikbud.go.id/detail_peraturan?main=3104 [Accessed: Oct. 16, 2023].
3. B. Trilling, C. Fadel, 21st Century Skills: Learning for Life in Our Times, John Wiley & Sons, 2009.
4. Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 20 of 2018 concerning Strengthening Character Education in Formal Education. [Online]. Available: https://peraturan.go.id/id/permendikbud-no-20-tahun-2018 [Accessed: Oct. 16, 2023].
5. D. Salirawati, Identifikasi problematika evaluasi pendidikan karakter di sekolah, Jurnal Sains dan Edukasi Sains 4(1) (2021). DOI: https://doi.org/10.24246/juses.v4i1p17-27
6. U. N. Sun'an, Supriyadi, S. Ridlo, W. Lestari, Literature review: Pentingnya pengembangan instrumen penilaian karakter peserta didik di era globalisasi, Proceedings of National Seminar Pascasarjana UNS. ISSN 26866404, 2023.
7. P. Black, D. Wiliam, Classroom Assessment and Pedagogy, Assessment in Education: Principles, Policy & Practice, 25(6) (2018) 551-575. DOI: https://doi.org/10.1080/0969594X.2018.1441807
8. S. C. Wong, Competency Definitions, Development and Assessment: A Brief Review, International Journal of Academic Research in Progressive Education and Development, 9(3) (2020) 95-114. DOI: http://dx.doi.org/10.6007/IJARPED/v9-i3/8223

9. L. W. Anderson & D. R. Krathwohl. A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition, Longman, 2001.
10. R. Martyasari, H. Suswanto, Sukarnati, The Mastery of Competencies the Skills Students of SMK Reviewed from The Talent, Creativity and Intrinsic Motivation of Students, Proceedings of The International Mechanical Engineering and Engineering Education Conferences (IMEEEC), 2016. DOI: https://doi.org/10.1063/1.4965803
11. T Lickona, Mendidik untuk Membentuk Karakter: Bagaimana Sekolah dapat Mengajarkan Sikap Hormat dan Tanggung Jawab, Bumi Aksara, 2019.
12. R. B. Kozma, M. Roth, Foreword, in: P. Griffin, E. Care (Eds.), Assessment and Teaching of 21st Century Skills, Springer, 2012.
13. P. Griffin, E. Care (Eds.), Assessment and Teaching of 21st Century Skills, Springer, 2012.
14. T. T. York, C. Gibson, S. Rankin, Defining and Measuring Academic Success, Practical Assessment, Research, and Evaluation, 20(5) (2019). DOI: https://doi.org/10.7275/hz5x-tx03
15. Indonesian Educational Standards, Curriculum and Assessment Agency, Assessment and Learning Guide. 2022. [Online]. Available: https://kurikulum.kemdikbud.go.id/rujukan [Accessed: Oct. 16, 2023].
16. R. J. Marzano, T. Heflebower, Teaching & Assessing 21st Century Skills, Marzano Research Laboratory, 2012.
17. L. Greenstein, Assessing 21st Century Skills: A Guide to Evaluating Mastery and Authentic Learning, Corwin, 2012.
18. G. Wiggins, The Case for Authentic Assessment, Practical Assessment, Research, and Evaluation, 2(1) (2019). DOI: https://doi.org/10.7275/ffb1-mm19.