Jurnal Penelitian Pembelajaran Fisika
Vol. 16 Issue 4 – October 2025, p407-414
p-ISSN 2086-2407, e-ISSN 2549-886X
Available Online at http://journal2.upgris.ac.id/index.php/JP2F
DOI: 10.26877/jp2f.v16i4.2848

Development of A Multidimensional Assessment Instrument on the Topic of Torque

Lola Ananda[1,2], Pujianto[1] and Edi Istiyono[1]
[1]Master Physics Education Department, Faculty of Mathematics and Natural Science, Universitas Negeri Yogyakarta. Jl. Colombo No. 1, Yogyakarta, Indonesia

[2]E-mail: lolaananda.2024@student.uny.ac.id

Received: 26 September 2025. Accepted: 28 October 2025. Published:30 October 2025.

Abstract. This research focuses on the development of a multidimensional assessment instrument aimed at evaluating high school students' learning outcomes in the topic of torque within physics education. The instrument integrates cognitive, affective, and psychomotor domains through a combination of multiple-choice questions for conceptual understanding and Likert-scale questionnaires for assessing students' interest and communication skills. The development process followed a systematic model consisting of planning, instrument construction, and pilot testing. The trial was conducted with 22 eleventh-grade students from SMAN 1 Dlingo, Bantul, Indonesia, selected via random sampling. Content validity was assessed using Aiken's V index, while internal consistency reliability was measured using Cronbach's Alpha with SPSS version 26.0. The results indicated that all items demonstrated high content validity ($V > 0.87$) and acceptable to excellent reliability coefficients: 0.74 for cognitive, 0.81 for affective, and 0.93 for psychomotor components. These findings affirm that the developed instrument is both valid and reliable, making it suitable for comprehensive evaluation of student learning outcomes in physics. Practically, the instrument enables teachers to diagnose students' conceptual mastery, motivation, and communication skills in an integrated manner, especially for abstract topics like torque. The novelty of this research lies in its multidomain assessment design tailored to the challenges of physics learning in the Indonesian high school context, supporting differentiated instruction and holistic evaluation aligned with the curriculum.

Keywords: instrument development, multidimensional assessment, torque

1. Introduction

Education serves as a pivotal foundation for individuals to cultivate their inherent potential. It is a critical necessity for every human being and functions as a primary indicator in advancing the intellectual development of a nation [1]. The progress of an era is largely determined by the quality of its education. Deliberate and systematic efforts to enhance educational quality are reflected in the design of instructional strategies and assessment systems that align with the characteristics and needs of learners [2]. The extent to which classroom learning objectives are achieved can be evaluated through students' learning outcomes. A key component of the instructional process is the evaluation system, commonly referred to as diagnostic assessment, which encompasses both cognitive and non-cognitive domains. As explained by Lee and Sawaki [3], cognitive diagnostic procedures aim to identify students' strengths and weaknesses in relation to their knowledge structures and processing skills, thereby serving as a reference for cognitive diagnostic assessment. Conversely, non-cognitive assessment seeks to determine students' emotional and psychological readiness to engage with learning materials [4]. Physics is widely recognized as a challenging subject due to its abstract nature and reliance on mathematical reasoning. These characteristics often lead to diminished student interest and conceptual misunderstanding. Educators must address these challenges to ensure the successful attainment of physics learning objectives in the classroom.

In physics education, several key aspects must be fulfilled to ensure students can comprehend the subject matter effectively, one of which is conceptual understanding. Among the various topics in physics, torque is considered particularly challenging, as it is difficult to comprehensively interpret students' learning achievement in this domain. A study conducted by [5] revealed that 73.68% of students identified torque as one of the most difficult topics in physics. This difficulty may be attributed to the instructional instruments employed by educators, which tend to be monotonous or lack sufficient variation. In response to this issue, it is imperative to develop more effective instructional instruments. One commonly used format is the multiple-choice test. This objective-type assessment is frequently utilized by educators to evaluate the extent of students' conceptual mastery, as it minimizes interpretive bias compared to essay-based assessments. Consistent with the work of [6] in their book *Quantitative and Qualitative Research Methods: Theory and Practice*, objective tests such as multiple-choice items are widely adopted in quantitative research due to their ability to yield measurable data that can be statistically analyzed to assess learning outcomes objectively. Moreover, multiple-choice tests are considered efficient, as they allow for the evaluation of numerous aspects within a limited time frame, thereby facilitating easier analysis. These items can also be enhanced by incorporating open-ended reasoning prompts for each selected answer, which further sharpens students' conceptual understanding.

The measurement of student learning outcomes is inherently complex and cannot be solely assessed through cognitive domains. A comprehensive evaluation must also encompass affective and psychomotor dimensions. As stipulated in the Indonesian Ministry of Education and Culture Regulation No. 66 of 2013 concerning assessment standards, the techniques and instruments employed in the 2013 curriculum include the evaluation of attitudes, knowledge, and skills competencies. This implies that all three domains must be developed in a balanced and integrated manner. [7] highlight that student's cognitive-affective profiles significantly influence the success of inquiry-based learning. This emphasis on affective integration is further supported by [8] through the CASTLE theory, which demonstrates that social and emotional interactions in digital learning environments contribute to deepper cognitive processing. [9] affirm that affectivity is an integral component of all human learning processes, suggesting that measuring student's attitudes and perceptions through Likert-scale instrument provides a more holistic view their learning readiness. This theoretical perspective reinforces the need to integrate affective and psychomotor indicators into physics assessment instruments.

Within the affective domain, students' interest in learning plays a pivotal role in achieving instructional goals. [10] emphasizes that students often encounter significant barriers in understanding physics due to the perception that the subject is inherently difficult to interpret and lacks appeal. When students initially perceive physics as uninteresting and tedious, classroom interaction between teachers and students tends to diminish, thereby impeding the learning process. This phenomenon is closely linked to the psychomotor domain, particularly communication skills. In physics education, communication is a critical component that supports students' learning outcomes—especially during laboratory activities. The ability of students to articulate experimental results through presentations and classroom discussions is essential. This aligns with the framework of 21st-century skills, commonly referred to as the 6C: Critical Thinking, Creativity, Collaboration, Communication, Character, and Citizenship [11]. These competencies are considered fundamental in preparing students to navigate global challenges and the complexities of life in the digital age [12]. Therefore, enhancing students' communication skills is imperative not only to foster greater interest in learning physics but also to enrich classroom dynamics and create a more engaging and vibrant educational experience.

In the cognitive domain, multiple-choice assessment procedures can be aligned with the levels of taxonomy employed. The interpretation of scores across taxonomic levels varies depending on item difficulty. Scoring within the cognitive domain may also be conducted using a dichotomous system, where correct responses are assigned a value of 1 and incorrect responses a value of 0. In the psychomotor domain, one commonly used assessment method is the Likert scale. This scale is designed to measure attitudes, opinions, and perceptions through a checklist format. A study by [13] demonstrated how students' attitudes can be assessed using a non-test instrument in the form of a Likert-based attitude questionnaire. The response options included: Always, Often, Sometimes, Rarely, and Never , structured according to an instrument blueprint in checklist format. Any instrument developed for educational assessment must undergo validity and reliability testing. These psychometric evaluations are essential to ensure that the instrument produces accurate and trustworthy data [14]. Once an instrument is confirmed to be valid and reliable, it is deemed appropriate for use, allowing researchers to apply it with confidence. Research conducted by [15] highlights the utility of the Quest software program as a data analysis tool for empirically determining item validity, reliability, and difficulty levels in instrument development.

Based on the urgency above, researchers need to develop instruments that cover all domains (cognitive, affective, and psychomotor) so that learning objectives can be achieved optimally and comprehensively. Torsion material is still a material that causes students to experience misconceptions due to its abstract nature and a lack of student interest in learning physics, which is often perceived as a difficult and intimidating subject. Research conducted by [16] revealed that students perceive torsion as one of the most difficult topics in physics, indicating the need for understanding. Consistently, identified misconceptions using the Rotational and Rolling Motion Concept Survey (RRMCS), revealing that 58,30 % of high school students experienced misconceptions and 38,10% did not understand the concept [17]. These findings indicate a gap in existing assessment practices, which often focus solely on cognitive outcamas and overlook affective and psychomotor dimensions. Therefore, Teachers need to foster curiosity so that students begin to actively engage in questioning and gradually develop a strong conceptual understanding of the material, supported by empirically tested multidimensional assessment instruments.

2. Method

2.1. *Research Design*

This study was designed to develop an assessment instrument on the topic of torque, aimed at measuring senior high school students' conceptual understanding, interest in learning, and communication skills. The development framework adopted in this research refers to both test and non-test instruments, as outlined by [14]. The process involved three primary stages of modification: planning, instrument trial, and measurement. However, the present study was limited to the instrument trial stage.

In the planning stage, indicators were derived from curriculum objectives and relevant literature, followed by the construction of test blueprints and item development for both multiple-choice and likert-scale formats. The construction of the non-test instrument was integrated with the prosedural steps of test instrument development, while maintaining adherence to the methodological principles specific to non-test instrument design. This integration ensured that the resulting instrument aligns with both cognitive and affective assessment standards in physics education.

To ensure the validity and reliability of the instrument, the trial stage was implemented with 22 eleventh-grade students selected through random sampling. Content validity was evaluated by expert judgment using Aiken's V indeks, while internal consitency reliability was measured using Cronbach's Alpha via SPSS version 26.0. The trial data were also used to conduct item analysis and descriptive statistics, allowing refinement of the instrument based on empirical evidence. This structured implementation of the trial stage provide a strong foundation for confirming the instrument's feasibility and psychometric quality prior to broader application.

2.2. *Research Sample*

The subjects involved in the instrument trial consisted of 22 eleventh-grade students during the second semester at SMAN 1 Dlingo, Bantul. The sampling technique employed in this study was random sampling.

2.3. *Research Instrument*

The instrument developed in this study comprised 10 multiple-choice items assessing conceptual understanding, 20 questionnaire items measuring students' interest in learning, and 14 questionnaire items evaluating communication skills, along with a validation sheet. The focus of the instrument development was on the torque material for eleventh-grade students, based on the *Kurikulum Merdeka* framework.

2.4. *Data Analysis Technique*

Content validity analysis was conducted to evaluate the extent to which each item in the instrument accurately represents all aspects of the intended construct, based on expert judgment. The validation process employed Aiken's V statistic to determine the appropriateness and relevance of the developed instrument. The data, in the form of expert rating scores, were systematically analyzed using Aiken's V formula, which is expressed as follows:

$$V = \frac{\Sigma s}{n(c-1)} \quad (1)$$

where V as The overall content validity coefficient is interpreted based on the value of V, which determines whether the instrument demonstrates acceptable content validity. According to Aiken [14], the interpretation of Aiken's V index can be referenced in Table 1.

Table 1. Category of Aiken's V index.

Aiken's V index	Validity
$0 < V \leq 0.4$	Low
$0.4 < V \leq 0.8$	Medium
$0.8 < V \leq 1$	High

The data analysis of the instrument trials in this study involved two tests: Item validity and Reliability. The calculation of item validity was performed on all instruments, namely the instruments for concept understanding, interest, and communication skills. The validation was conducted using the SPSS version 26.0 application. The approach used was internal consistency, which involved testing the correlation between item scores and the total score. For the formative assessment concept understanding instrument, a biserial correlation was used. The reliability test was conducted on all instruments: the instruments for concept understanding, interest, and communication skills. Cronbach's Alpha was the analysis technique used for this test with the SPSS version 26.0 application.

3. Result and Discussion

Step-by-step guidelines served as the reference used to develop this research instrument, which was adapted from the [18] model. The steps taken in the assessment instrument development process are as follows:

3.1. *Instrument Design Stage*

This research is a developmental research aimed at producing an assessment instrument for students' conceptual understanding, interests, and communication skills that can be used in physics learning on torque. The instrument developed consists of multiple-choice questions (cognitive) and a questionnaire to assess students' interests (affective) and communication skills (psychomotor). The assessment instrument outline is designed as a guideline for designing test items, taking into account learning outcomes, teaching materials, question formats, and indicators of interest and communication skills to be measured. This development aims to ensure that each test item aligns with the competencies targeted in the independent curriculum.

The cognitive ability aspects and indicators, namely conceptual understanding developed in this instrument, refer to the conceptual understanding indicators according to [19], namely: a) translation, b) interpretation, c) extrapolation. It is further explained that translation uses operational verbs such as translating, changing, defining, illustrating, and re-explaining. Meanwhile, interpretation uses operational verbs such as interpreting, explaining, differentiating, and describing. And finally, extrapolation is drawing conclusions from something that is already known. Indicators from [20] are also adopted, including interpretation, exemplifying, classifying, generalizing, inferring, comparing, and explaining.

The aspects and indicators of communication skills developed in this instrument refer to [21], which include: a) Ability to convey ideas and opinions verbally, b) Ability to listen and respond to the other person, c) Clarity in pronunciation and sentence structure, d) Politeness and ethics in communication, e). Use of body language and facial expressions, f). Confidence in speaking. In addition, indicators from [22] were also adopted, including: a) Ability to ask questions, b) Ability to express opinions, c) Ability to respond to information, d) Ability to argue, e) Ability to listen and respect the opinions of others, f) Politeness in speaking, g) Confidence in speaking, h) Critical and tolerant attitude towards differences of opinion.

Based on the results of the synthesis, the aspects and indicators of learning interest used refer to [19], which include: a) Interest in physics material, b) Desire to know more deeply, c) Feelings of pleasure when taking lessons, d) Active involvement in the learning process, e) Enthusiasm in completing physics assignments. These indicators were operationalizes into 20 questionnaire item to assess learning interest, 14 items to assess conceptual understanding. Scoring for the multiple-choice items was binary, assigning 1 point for correct answers and 0 for incorrect ones. The questionnaire items employed a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree", allowing for standardized scoring across affective and psychomotor domains. The use of Likert scales in educational research is widely supported due to their ability to capture nuanced attitudes and behaviors, and their psychometric robustness when properly constructed and validated [23].

To ensure content validity, expert judgment was analyzed using Aiken's V index. All items across the three domains demonstrated high content validity, with values ranging from 0.87 to 1.00. the average V scores were 0.93 (cognitive), 0.91 (affective), and 0.95 (psychomotor), indicating high expert agreement across all domains. Item-total correlation analysis showed that all cognitive items had significant biserial coefficients (0.42-0.86), while affective and psychomotor items had pearson coeffients between 0,45 and 0,78. Reliability coefficients were 0.74 (cognitive), 0.81 (affective), and 0.93 (psychomotor), confirming strong internal consistency. All items met validity and reliability standards and were retained.

The developed instrument items need to be validated to ensure their suitability. Validation is carried out by two practical validators, both high school physics teachers, who will assess and provide input and suggestions for improvements to the developed instrument. The validation results are then analyzed to obtain a Sig (2-tailed) value and Pearson correlation using point biserial with the help of SPSS 26.0, as shown in the table 2 and indeks V aiken as shown in the tabel 3 and 4.

Table 2. Result of multiple choice questions.

Number	*Sig (2-tailed)*	*Pearson Correlation*	Category	Interpretation
1	0.000	0.862	Valid	Very high
2	0.000	0.747	Valid	High
3	0.009	0.541	Valid	Medium
4	0.000	0.806	Valid	Very High
5	0.001	0.680	Valid	High
6	0.000	0.747	Valid	High
7	0.009	0.541	Valid	Medium
8	0.000	0.862	Valid	Very High
9	0.000	0.684	Valid	High
10	0.000	0.806	Valid	Very high

Table 3. Result of V Aiken communication.

Number	V value	Interpretation
3,4,5,9,10	0.95	High
1,7,11,13,14	0.93	High
6,8,12	0.91	High
2	0.87	High

Table 4. Result of V Aiken interest.

Number	V value	Interpretation
3,5,7,9,11,17,19,20	1	High
4,6,10,13,14,15,18	0.97	High
1,16	0.95	High
2,8	0.93	High
12	0.91	High

Tables 3 and 4 show the results of the Aiken V-index for 14 communication skills questionnaire items and 20 physics learning interest questionnaire items in the high category. Therefore, the developed assessment instrument is declared highly valid. Several suggestions and input from the validator were needed for improvement, such as providing instructions or command sentences in the instrument, and using the teacher observation sheet for validation assessment. After improvements were made in accordance with the suggestions and input, the assessment instrument was declared suitable for trial testing.

3.2. *Instrument Testing Phase*

This study was conducted using random sampling techniques with 11th-grade students at SMA Dlingo, Bantul. The trial subjects were 22 students, based on the number of students present at the time. The instrument trial data were then analyzed to determine validity and reliability for each domain.

3.3. *Validity of Assesment Instruments*

Based on the calculation results using SPSS and V Aiken in tables 2,3, and 4, the results show that the instrument items for each domain are declared valid. In the conceptual understanding ability instrument, item 2, questions numbered 3 and 7, received a sufficient level of validation, while 4 questions numbered 1, 4, 8, and 10 received a very high level of validation, and the rest received a high level of validation. In communication skills and interests, the validation results were the same, namely high.

3.4. *Reliability of Assesment Instruments*

Based on the reliability analysis, the comprehension ability instrument yielded a Cronbach's alpha value of 0,74, which falls within the acceptable range of 0,70 to 0,79 and is considered sufficiently reliable for classroom-level assessments [24]. The instrument consisted of 10 multiple-choice items, all contributing positively to internal consistency. For communication skills, the realiability coefficient was 0,93, indicating excellent reliability as it exceeds the 0,90 threshold commonly cited for behavioral constructs [25]. Meanwhile,the learning interest instrument produced a reliability score of 0,81, categorized as good reliability within 0.80 to 0.89 range. These results demonstrate that all three instruments (cognitive, affective, and psychomotor) are psychometrically sound and suitable for evaluating multidimensional learning outcomes in physics education.
The reliability data can be seen in Table 5.

Table 5. Result of each indikator multidimensional.

Indicator	Communication	Interest	Conceptual Understanding
Item variants	8,46	2,44	Cronbach's Alpha
Total of Variants	103,64	14,55	
Item Reliability	0,93	0,81	0,74

4. Conclusion

This research produced an assessment instrument designed to measure three important aspects of physics learning on torsion: conceptual understanding (cognitive), communication skills (psychomotor), and interest in learning physics (affective). The instrument was developed through planning, developing a grid, expert validation, and field testing with 22 eleventh-grade high school students. Content validation using Aiken's V-index showed that all instrument items were in the high validity category ($V > 0.87$). Item validity testing using SPSS demonstrated that all items were valid, with Pearson correlations ranging from moderate to very high. Reliability testing using Cronbach's Alpha demonstrated adequate reliability: 0.74 for conceptual understanding, 0.81 for interest in learning, and 0.93 for communication skills. These findings demonstrate that the developed instrument not only meets validity and reliability criteria but also accommodates the needs for comprehensive physics learning assessment in accordance with the Independent Curriculum approach. Thus, this instrument is suitable for use by educators as an objective and multidimensional evaluation tool in measuring student learning outcomes in torque material. However, this study is limited by its sample scope and contextual specificity, which may affect the generalizability of the instrument across diverse educational settings. Future research is recommended to expand validation efforts across broader populations, explore longitudinal impacts on student learning, and adapt the instrument for other physics topics or interdisciplinary contexts to enhance its applicability and scalability.

References

[1] Nurjannah H, Saputro A, Maddatuang M and Fikri M J N 2020 Implementation of the Treffinger Learning Model in Geography Instruction *LaGeografia* **19** 113–127

[2] Sufyadi S 2021 Learning and Assessment Guidelines for Elementary and Secondary Education Levels, Assessment and Learning Center, Research and Development and Book Agency, Ministry of Education, Culture, Research, and Technology, 1.

[3] Lee Y W and Sawaki Y 2009 Cognitive diagnosis approaches to language assessment: an overview *Lang. Assess. Q*. **6** 172–189

[4] Hati S M 2021 The Effectiveness of Using the Quizizz Application in Conducting Non-Cognitive Diagnostic Assessments of Grade 12 Social Studies Students with Cross-Interests at YPHB Senior High School, Bogor City *Arus Journal of Education* **1** 70

[5] Aulia M, Normilawati, Yuliani H and Azizah N 2021 Needs Analysis of Contextual-Based Comic Learning Media for Torque Material in Senior High School *J. Sci. Educ. Phys.–COMPTON* **8** 72–81

[6] Permana S A A, Rohman M M, Arta D N C, Bani M D and Bani G A 2023 *Quantitative and Qualitative Research Methods: Theory and Practice* (Yogyakarta: GET Press Indonesia) ISBN 978-623-198-668-9 ed Sulung N

[7] Feng S, Lu C, Shen S, Ning N, Wang M, Wang S 2025 Identifying students' cognitive-affective profiles and associations with gender and behaviors in virtual scientific inquiry *Educ. Technol. Res. Dev.* **73**

[8] Schneider S, Beege M, Nebel S, Schnaubert L, Rey GD 2022 The Cognitive-Affective-Social Theory Of Learning In Digital Environments (CASTLE) *Educational Psychology Review* **34** 1-38

[9] Fossa and Cortés-Rivera C 2023 Affectivity and Learning: Bridging the Gap Between Neurosciences, Cultural and Cognitive Psychology *Springer*

[10] Saftari M and Fajriah N 2019 Assessment of the Affective Domain through Attitude Scale Evaluation to Measure Learning Outcomes *Edutainment: Journal of Educational Science and Pedagogy* **7**

[11] Van der Voet and B. Steijn 2021 Team innovation through collaboration: How visionary leadership spurs innovation via team cohesion *Public Manage. Rev.* **23** 1275–1294

[12] Yusup F 2018 Validity and Reliability Testing of Quantitative Research Instruments *Tarbiyah: Journal of Educational Science* **7**

[13] Kusumawati A E, Supahar and Pebriana I N 2024 Development of Assessment Instruments on Global Warming to Measure Students' Science Process and Collaboration Skills *Diffraction: Journal for Physics Education and Applied Physics* **6**

[14] Istiyono E 2020 *Development of Assessment Instruments and Analysis of Physics Learning Outcomes Using Classical and Modern Test Theories* (Yogyakarta: UNY Press)

[15] Aiken L R 1985 Three coefficients for analyzing the reliability and validity of ratings *Educ. Psychol*. Meas. **45** 131–142

[16] Normilawati, Aulia M, Yuliani H and Azizah N 2021 Description of the Requirement For Contextual-Based Comic Learning Media On Torque Material In High School *Jurnal Ilmiah Pendidikan Fisika-COMPTON* **8** 72–81

[17] Rafika R and Syuhendri S 2021 *Students' Misconceptions On Rotational And Rolling Motions* Journal of Physics: Conference Series

[18] Oriondo L L and Antonio E M 1998 *Evaluating Educational Outcomes: Tests, Measurement and Evaluation* (Quezon City: Rex Book Store)

[19] Puri P R A and Perdana R 2023 Analysis of Senior High School Students' Conceptual Understanding of Static Fluid and Its Improvement through the Visualization Auditory Kinesthetic Learning Model *Magneton: Journal of Innovation in Physics Learning* **1** 93–101

[20] Gultom C, Sihombing R and Harahap S H 2024 *Evaluation of Oral Communication Proficiency in Indonesian Language Learning* (Medan: Universitas Negeri Medan)

[21] Marfuah 2017 Enhancing Students' Communication Skills through the Jigsaw-Type Cooperative Learning Model *JPIS: Journal of Social Science Education* **26** 148–160

[22] Nurdanti V Y 2020 The Influence of Learning Interest on Students' Physics Learning Outcomes *Schrödinger: Journal of Physics Education* **1** 101–106

[23] Mohd Rokeman 2024 Likert Measurement Scale in Education and Social Sciences: Explored and Explained *EDUCATUM Journal of Social Sciences* **10** 84–94

[24] Nunnally and Bernstein I H 1944 *Psychometric Theory* 3rd ed. (New York, NY, USA: McGraw-Hill)

[25] Barbera M M, VandenPlas and Stains J M 2021 Reliability and construct validity of the affective dimension of the Colorado Learning Attitudes about Science Survey for Chemistry," J. Chem. **98** 519–529