Volume 17
Number 2
Dec 2025

Journal of
The Indonesian Society of Integrated Chemistry
Published by Jambi University, Chemistry Education Study Program

P-ISSN 2085-1715 & E-ISSN 2621-5543

JISIC

Development of an Assessment Instrument Based on Wiggins and McTighe's Six Facets of Understanding for Chemical Equilibrium

Legia Ladjesta Dwiraga[1*], Betty Holiwarni[2], Putri Adita Wulandari[3]

[1,2,3] Universitas Riau, Indonesia

ABSTRACT

This study aimed to develop an assessment instrument grounded in Wiggins and McTighe's Six Facets of Understanding and to examine its feasibility through expert validation and classical item analysis. The study employed a research and development (R&D) approach using the ADDIE model (Analyze, Design, Develop, Implement, Evaluate). The development process in this study was conducted up to the Development stage. Expert validation (three validators) indicated that the instrument was highly valid in terms of construction (94%) and achieved perfect validity for content/material and language aspects (100%). A small-group trial also showed positive user responses, with high readability (91%) and adequate time allocation (74%). Teachers' responses indicated very high ratings for relevance/fit (100%), language clarity (87.5%), and usefulness (95.83%). Item analysis showed that all 10 items were valid, and the instrument demonstrated high reliability (Cronbach's alpha = 0.846). The difficulty indices were moderate, and most items showed acceptable-to-good discrimination. Overall, the developed instrument is feasible to proceed to the Implementation stage in future research.

Keywords: alternative assessment; learning outcomes; Merdeka Curriculum (Independent Curriculum); chemical equilibrium.

Received: Juny, 11 2025;	* Corresponding author: legialadjesta2003@gmail.com
Revised: August, 29 2025;	DOI: https://doi.org/10.22437/jisic.v17i2.45247
Accepted: September, 10 2025	

INTRODUCTION

The Merdeka Curriculum has been officially established and is currently used by Indonesian educational institutions. Its implementation calls for instruction that prioritizes students' character development in line with the Pancasila Student Profile, as well as flexible learning that responds to students' competency-development needs and focuses on essential content so that students are adequately supported to engage in deep and meaningful learning (Barlian et al., 2022). This curricular orientation toward deep and meaningful learning ultimately aims to ensure the achievement of learning objectives and optimal student outcomes. Therefore, designing appropriate assessment instruments is one of the key essentials for measuring the extent to which these learning objectives are attained.

In the Merdeka Curriculum, learning objectives are derived from learning

outcomes. In general, learning objectives are formulated using Bloom's Taxonomy, which hierarchically classifies cognitive abilities from the most basic to the most advanced levels. However, according to the National Curriculum and Assessment Standards Agency (BSKAP), current chemistry learning outcomes emphasize students' understanding (BSKAP, 2024). Meanwhile, "understanding" in Bloom's Taxonomy is typically positioned at the C2 level, which is considered a lower-order cognitive level and thus less aligned with the demands of the current Merdeka Curriculum. In response, the BSKAP (2024) Learning and Assessment Guidelines recommend alternatives to Bloom's Taxonomy for formulating learning objectives, including Wiggins and McTighe's framework and Marzano's taxonomy. Wiggins and McTighe's formulation—through the Six Facets of Understanding—focuses on deep understanding and is therefore relevant to the current Merdeka Curriculum. Using this framework, teachers can assess students' understanding and competence through the facets of explanation, interpretation, application, perspective, empathy, and self-knowledge. These facets position students not only to know the content, but also to grasp facts and meaning by transferring and applying their knowledge and skills to new situations based on their learning experiences (Elisa, E., 2022; Wiggins & McTighe, 2005).

METHODS

The method used in this study was Research and Development (R&D). The development procedure followed the ADDIE model, which consists of five stages. In the first stage (Analyze), the researchers conducted a needs analysis and reviewed supporting data sources to develop the instrument. After the required data were collected, a draft product was designed in the second stage (Design), which involved determining the item format, developing validation sheets, constructing the Six Facets of Understanding–based assessment instrument, and preparing the scoring guidelines. Once the product draft had been completed, the study proceeded to the third stage (Development). Prior to the tryout, the developed Six Facets of Understanding–based assessment instrument was validated by expert validators to evaluate its content feasibility. The product feasibility criteria are presented in Table 1.

Table 1. Validation Feasibility Criteria

Percentage	Validation Criteria
80.00–100.00	Valid
60.00–79.99	Sufficiently valid
50.00–59.99	Less valid
0–49.99	Not valid

(Riduwan, 2012)

After the instrument was deemed feasible, it was piloted in a small-group trial to obtain user feedback and subsequently followed by an item analysis to measure:

Validity analysis.

To determine the validity of the assessment items, the Pearson product–moment correlation formula was used, as follows:

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\{N\sum X^2 - (\sum X)^2\}\{N\sum Y^2 - (\sum Y)^2\}}} \text{.....(1)}$$

Notes:

r = correlation coefficient between variables X and Y
N = number of respondents
X = score for an individual item
Y = total score

This analysis can be conducted using SPSS software. After the item validity coefficients are obtained, the results are compared with the critical r value from the correlation table at the 5% and 1% significance levels with $df = N - 2$. If r (calculated) > r (table), the item is considered valid at the selected significance level.

Reliability Analysis

Reliability testing is conducted to determine whether the items consistently measure the intended construct. The reliability coefficient in this study was calculated using Cronbach's alpha (Sugiyono, 2019).

$$r_{11} = \left[\frac{k}{k-1}\right]\left[1 - \frac{\sum \sigma_b^2}{\sigma_t^2}\right] \quad \text{.....(2)}$$

r_{11} = instrument reliability coefficient
k = number of items
$\sum\sigma_i^2$ = sum of the item-score variances (for item i)
σ_t^2 = total score variance

The r value obtained from the Cronbach's alpha calculation was then compared with the critical r value from the Pearson product–moment correlation table at $\alpha = 0.05$ with $df = N - 2$ (N = number of students).

Item Difficulty Analysis

Item difficulty reflects the proportion of students who answer an item correctly. By examining the difficulty index, researchers can identify items of acceptable quality that are suitable for use. The difficulty index was calculated using the following formula (Sumarna, 2009).

$$P = \frac{\sum X}{SmN} \text{ ..... (3)}$$

Notes:

p = difficulty indeks
$\sum X$ = total item score
S_m = maximum possible score
N = number of students

The criteria for item difficulty are presented in Table 2.

Table 2. Item Difficulty Criteria

Difficulty Index	Item Category
0.00–0.30	Difficult
0.31–0.70	Moderate
0.71–1.00	Easy

(Sudjana, 2011).

Item Discrimination Analysis

Item discrimination examines how well an item differentiates among students with high, moderate, and low ability levels (Ida Ayu Gde Yadnyawati, 2019). Discrimination was determined by ranking all test-takers by total score from highest to lowest. The participants were then divided into the upper 27% (high total scores) and the lower 27% (low total scores) groups (Sumarna, 2009). The discrimination index

was calculated using the following formula (Sukardi, 2015):

$$DP = \frac{BA - BB}{\frac{1}{2}.N} \text{ ..... (4)}$$

Notes:

BA = upper group (high-performing students)
BB = lower group (low-performing students)
N = number of test-takers

The classification of item discrimination is shown in Table 3.

Table 3. Item Discrimination Criteria

Discrimination Index	Criteria
< 0.00 (negative)	Very poor
0.00–0.20	Poor
0.21–0.40	Fair
0.41–0.70	Good
0.71–1.00	Very good

The final output of this development study was a Six Facets of Understanding–based assessment instrument grounded in Wiggins and McTighe's framework that is feasible for large-scale testing at the Implementation stage. However, the Implementation stage was not conducted in this study; the development process was carried out only up to the Development stage. The Evaluation stage was implemented formatively, meaning that evaluation was performed at each stage of the ADDIE model.

The data sources in this study—including the small-group trial and item analysis were Grade XI students from State Senior High School 4 Pekanbaru (SMA Negeri 4 Pekanbaru) and State Senior High School Plus of Riau Province (SMA Negeri Plus Provinsi Riau). Data were collected using validation sheets, the Six Facets of Understanding–based assessment items, and a user response questionnaire. The validation sheets were used to determine the validity of the assessment items in terms of construction, content/material, and language through review by three experts. The assessment items were used to obtain students' scores and to conduct item analysis. The user response questionnaire was administered to collect feedback from the teacher and students regarding the developed items.

RESULTS AND DISCUSSION

Understanding enables students not only to know the content, but also to grasp facts and meaning by being able to transfer, use, and apply their knowledge or skills to new situations based on their learning experiences (Wiggins & McTighe, 2005). Therefore, the ability to transfer what has been learned is essential for students to internalize knowledge effectively and achieve meaningful learning. Wiggins and McTighe further developed the **Six Facets of Understanding**, indicating that students should demonstrate competence across all six facets to determine the depth of understanding attained after instruction. This orientation toward deep understanding aligns with the current Merdeka Curriculum and is also emphasized in the Merdeka Curriculum Learning and Assessment Guidelines (BSKAP, 2024).

Validation of the Wiggins and McTighe Six Facets–Based Assessment Instrument

The Six Facets of Understanding–based assessment instrument was developed for the topic of chemical equilibrium. Validation was conducted by three expert validators who reviewed the instrument in terms of construction, content/material, and

language. Each validator was asked to evaluate the developed items and provide suggestions for improvement, enabling the level of validity to be determined. The validation process involved two rounds of revision, and the validators provided final ratings at the end of the validation stage. The validation results are presented in Figure 1.

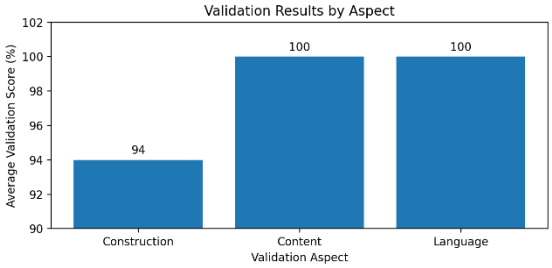

Figure 1. Average Validation Scores

Expert validation results indicated that the 10 developed items were **highly valid** in terms of construction, content/material, and language, and therefore suitable for the small-group trial.

The small-group trial involved 10 Grade XI students at **State Senior High School Plus of Riau Province**. Both the test-takers and the chemistry teacher completed a user response questionnaire to provide feedback on the developed instrument. The results of the teacher and student response analyses are presented in Tables 4 and 5.

Table 4. Teacher User Response Results

Aspect	Score (%)	Category
Relevance/Alignment	100	Very High
Language	87.5	Very High
Usefulness	95.83	Very High
Average (%)	**94.44**	**Very High**

Table 5. Student User Response Results

Aspect	Score (%)	Category
Item Readability	91	Very High
Time Adequacy	74	High
Average (%)	**82.5**	**Very High**

Overall, user feedback was positive regarding the developed assessment instrument. Next, to examine item quality, classical test theory–based item analysis was conducted, including tests of item validity, reliability, discrimination, and difficulty (Ida Ayu Gde Yadnyawati, 2019).

Table 6. Summary of Item Correlation Coefficients and Item Status

Item No.	r (calculated)	r (table)	Status
1.	0,720	0,374	Valid
2	0,488	0,374	Valid
3	0,735	0,374	Valid
4	0,708	0,374	Valid
5	0,710	0,374	Valid
6	0,733	0,374	Valid
7	0,625	0,374	Valid
8	0,694	0,374	Valid
9	0,492	0,374	Valid
10	0,627	0,374	Valid

Based on Table 6, all 10 developed items were deemed valid and suitable for use. In addition to validity testing, reliability analysis was conducted using Cronbach's alpha. The obtained Cronbach's alpha coefficient of 0.846 indicates that the 10 items in the Wiggins and McTighe Six Facets of Understanding–based assessment instrument demonstrate high reliability.

Characteristics of the Six Facets of Understanding–Based Assessment Instrument

The items that were confirmed to be valid and reliable were further examined through item characteristic analysis to determine their difficulty levels and discrimination indices. This item characteristic analysis was administered to 30 students at State Senior High School 4 Pekanbaru. The item difficulty results are presented in Table 7.

Table 7. Summary of Item Difficulty Levels

Item No.	Value	Category
1	0.70	Moderate
2	0.63	Moderate
3	0.68	Moderate
4	0.45	Moderate
5	0.45	Moderate
6	0.67	Moderate
7	0.70	Moderate
8	0.57	Moderate
9	0.57	Moderate
10	0.70	Moderate

The difficulty analysis indicated that the 10 developed items are acceptable and suitable for use. The obtained difficulty indices range from 0.30 to 0.70, indicating well-functioning items (Zainal, 2012). Each item was then examined using item discrimination analysis to evaluate its ability to differentiate among students with high, moderate, and low ability levels. The discrimination results are presented in Table 8.

Table 8. Summary of Item Discrimination Indices

Item No.	Discrimination Index	Category
1	0.6	Good
2	0.4	Fair
3	0.67	Good
4	0.53	Good
5	0.6	Good
6	0.53	Good
7	0.6	Good
8	0.3	Fair
9	0.4	Fair
10	0.6	Good

The obtained discrimination indices were all ≥ 0.30, indicating that the developed Six Facets of Understanding–based assessment instrument can adequately distinguish among students with high, moderate, and low ability levels (Ida Ayu Gde Yadnyawati, 2019).

Through a development procedure consisting of the following stages: (1) Analyze, (2) Design, and (3) Development—with formative evaluation conducted at each stage—the study produced an instrument aligned with Wiggins and McTighe's Six Facets of Understanding framework for deep understanding, consistent with the expectations of the Merdeka Curriculum. Expert reviewers also indicated that the developed instrument met very high quality standards across construction, content/material, and language aspects. This further suggests that the items were appropriately aligned with the indicators derived from the Six Facets of Understanding.

User responses likewise showed that the developed items can be effectively used in classroom instruction and can help teachers identify content areas that require greater emphasis. The items were also considered appropriate for senior high school (SMA) students, easy to read, and understandable.

Based on the findings above, it can be concluded that the developed Six Facets of Understanding–based assessment instrument is suitable for use, as it has demonstrated sound item characteristics.

.

CONCLUSION

The Six Facets of Understanding–based assessment instrument for Grade XI chemical equilibrium (or equivalent SMA/MA level) was developed using the ADDIE model through the Analyze, Design, and Development stages; the Implementation stage was not conducted because the study focused on instrument development. Expert validation indicated that the instrument was **valid** (content/material and language: **100%**; construction: **94%**). User feedback was also positive, with teacher ratings showing very high relevance/alignment (**100%**), strong language clarity (**87.5%**), and high usefulness (**95.83%**), while student responses indicated high readability (**91%**) and adequate time allocation (**74%**). Item analysis confirmed that all 10 items were valid and reliable, with acceptable difficulty levels and discrimination indices. Future studies should proceed to the **Implementation** stage to evaluate the instrument's effectiveness in real assessment settings and to adapt the Six Facets framework for assessment development in other chemistry topics.

REFERENCES

Barlian, U. C., Solekah, S., & Rahayu, P. (2022). Implementasi Kurikulum Merdeka dalam meningkatkan mutu pendidikan. *Journal of Educational and Language Research, 10*(1), 1–52. https://doi.org/10.21608/pshj.2022.250026

Badan Standar Kurikulum dan Asesmen Pendidikan (BSKAP). (2024). *Panduan pembelajaran dan asesmen pendidikan anak usia dini, pendidikan dasar, dan pendidikan menengah (Edisi revisi tahun 2024)* (pp. 1–72). BSKAP.

Elisa, E. (2022). Peningkatan keterampilan dasar laboratorium kimia teknik melalui praktikum mandiri berbantukan media laboratorium virtual di masa pandemi COVID-19. *Journal of The Indonesian Society of Integrated Chemistry, 14*(1).

Ridho, D., Juwitaningsih, T., Azzura, S., Saktiono, A., Rambe, A. K., Hutagaol, J. M., & Siregar, L. S. (2025). Perspektif mahasiswa matematika tentang pencampuran ekstrak lidah buaya (*Aloe vera*) dengan larutan asam lemah terhadap kesehatan dan kelembutan rambut. *Jurnal Riset Rumpun Ilmu Pendidikan, 1*(1), 589–600.

Ridho, D., Juwitaningsih, T., Prayudi, N. E. P., Syifa, N. M., Syahbilah, P., & Hasibuan, F. I. (2025). Perspektif mahasiswa jurusan matematika kelas kimia umum terhadap pemanfaatan bunga telang dan cuka sebagai tinta alami ramah lingkungan. *Jurnal Riset Rumpun Matematika Dan Ilmu Pengetahuan Alam, 4*(1), 149–159. https://doi.org/10.55606/jurrimipa.v4i1.5014

Ridho, D., Juwitaningsih, T., Samosir, R. A., Pratiwi, A. J., Ardila, M., Kurniawan, E. D. A., Sianipar, M. W. S., Alisya, S. N., Hutahuruk, N. B., & Tafonao, F. S. (2025). General chemistry class perspective on the utilization of young coconut water to treat fever symptom. *Hydrogen: Jurnal Kependidikan Kimia, 13*(3), 530–540. https://doi.org/10.33394/hjkk.v13i3.15819

Riduwan. (2012). *Metode dan teknik menyusun proposal penelitian*. Alfabeta.

Setyaedhi, H. S., Mustaji, & Fitri, C. (2023). Empirical quality of final exam questions in a learning management system-based course. *JPI (Jurnal Pendidikan Indonesia), 12*(1), 78–88. https://doi.org/10.23887/jpiundiksha.v12i1.52262

Siregar, A. D., Ravico, & Harahap, L. K. (2022). Upaya meningkatkan kreatifitas mahasiswa menggunakan bahan ajar kimia berbasis project based learning (PjBL). *CHEDS: Journal of Chemistry, Education, and Science, 6*(1), 1–6. https://doi.org/10.30743/cheds.v6i1.5500

Sudjana. (2011). *Penilaian hasil proses belajar mengajar*. PT Remaja Rosdakarya.

Sugiyono. (2019). *Statistika untuk penelitian*. Alfabeta.

Sugiyono. (2020). *Metode penelitian kuantitatif, kualitatif, dan R&D* (2nd ed.). Alfabeta.

Sukardi. (2015). *Evaluasi pendidikan: Prinsip dan operasionalnya*. PT Bumi Aksara.

Sumarna, S. (2009). *Analisis, validitas, reliabilitas, dan interpretasi*. PT Remaja Rosdakarya.

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). ASCD.

Yadnyawati, I. A. G. (2019). *Evaluasi pembelajaran* (I. K. Suda, Ed.). UNHI Press.

Zainal, A. (2012). *Evaluasi pembelajaran: Prinsip, teknik, prosedur*.