Jurnal Penelitian dan Evaluasi Pendidikan
Volume 29, No. 2, December 2025 (256-270)
Online: http://journal.uny.ac.id/index.php/jpep

Chemistry teachers' self-efficacy scale for differentiated instruction: Development and validation

Rais Assad Faiz*; Hari Sutrisno
Universitas Negeri Yogyakarta, Indonesia
*Corresponding author. E-mail: raisassad.2023@student.uny.ac.id

ARTICLE INFO

Article History
Submitted:
2 November 2025
Revised:
2 December 2025
Accepted:
24 December 2025

Keywords
chemistry teacher; differentiated instruction; instrument development; scale development; self-efficacy; teacher competence

Scan Me:

ABSTRACT

This research aims to develop and validate a new instrument for measuring chemistry teacher's self-efficacy for differentiated instruction. Data was collected using developed questions based on a literature study. The data were analyzed using exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and correlational techniques based on responses from 156 chemistry teachers in Yogyakarta who participated in an offline and online survey. The findings from EFA, CFA, and correlation analysis provide sufficient empirical evidence to support the convergent and discriminant validity of the instrument. The high Cronbach's coefficient alpha values show the instrument's good internal consistency and reliability. We documented a valid and reliable chemistry teachers' self-efficacy for differentiated instruction with six constructs consisting of 22 items by integrating the evidence from theory and data. The six-factor scale is named "efficacy in differentiating the learning process, efficacy in differentiating learning content, efficacy in differentiating learning products, efficacy in identifying student differences, efficacy in assessment, and efficacy in differentiating the learning environment". This instrument enriches the theoretical understanding of teacher self-efficacy in differentiated instruction by providing a domain-specific measurement model tailored for chemistry education. Practically, the scale offers teacher educators, policymakers, and professional development programs a diagnostic tool to identify teachers' strengths and areas needing support, enabling more targeted training interventions to enhance differentiated instruction in classroom practice.

This is an open access article under the **CC-BY-SA** license.

To cite this article (in APA style):
Faiz, R. A., & Sutrisno, H. (2025). Chemistry teachers' self-efficacy scale for differentiated instruction: Development and validation. *Jurnal Penelitian dan Evaluasi Pendidikan*, *29*(2), 256-270. https://doi.org/10.21831/pep.v29i2.91036

INTRODUCTION

Through Regulation No. 12 of 2024 of the Minister of Education, Culture, Research, and Technology of the Republic of Indonesia, the Indonesian Government formulated a new curriculum, the "*Kurikulum Merdeka*". One of the calls in the "*Kurikulum Merdeka*" is to implement differentiated instruction. Module 2.1 of the "*Guru Penggerak*" program states that teachers must implement differentiated instruction to accommodate students' diverse learning needs. The module emphasizes applying three differentiation strategies, namely (1) content differentiation, (2) process differentiation, and (3) product differentiation. The purpose of differentiated instruction, in general, is to carry out learning that emphasizes aspects of student learning interests, student readiness in learning, and student learning profiles (Marlina, 2019).

Differentiated instruction was first introduced by Carol Ann Tomlinson in 1999 as a form of criticism of *one-size-fits-all learning* that ignores the diversity of student needs (Fox & Hoffman, 2011; Subban, 2006). The diversity of students in the classroom is a necessity. This diversity includes differences in learning styles, motivations, abilities, needs, and interests (Suprayogi *et* *al.*, 2017). Differentiated instruction is considered to consider differences between students, recognize student strengths and accommodate student limitations (George, 2005; Heacox, 2012; Subban, 2006; Tomlinson *et al.*, 2003). Asriadi's (2023) meta-analysis of 63 related articles found that differentiated learning positively affects learning outcomes compared to traditional learning. These findings are supported by research (Cruz *et al.,* 2019; Whitley *et al.*, 2021), which states that by considering the unique characteristics and needs of each student, the application of differentiated instruction can increase student engagement, motivation, understanding, and learning skills, which ultimately leads to improved learning outcomes. In chemistry learning, a quantitative study by Gatune *et al.* (2022) reported that differentiated instruction significantly improved student achievement in chemistry learning.

Teachers are the ones most responsible for the success of implementing differentiated instruction in the classroom. However, implementing differentiated instruction is not as easy as imagined. Differentiated instruction is a complex teaching skill (Gaitas & Martins, 2017; Smets, 2017; van Geel *et al.*, 2019). Research conducted by the Dutch Inspectorate of Education (Inspectie van het Onderwijs, 2014; 2015a; 2015b) shows that teachers do not adapt their teaching to student differences. In addition, some teachers are not ready to face this task. Although teachers have received professional development training on differentiated instruction, they rarely apply it in the classroom because of the difficulty of implementing it and choose to apply traditional approaches (Dixon *et al.,* 2014; DeNeve *et al.*, 2015). Suprayogi *et al.*'s (2017) research revealed that although teachers have sufficient knowledge about differentiated instruction, the application of differentiated instruction is still limited.

The implementation of differentiated instruction that is not optimal is often caused by factors related to teachers, such as self-efficacy (Tobin & Tippett, 2014; Lavania & Nor, 2020). A study by Zelalem et al. (2022) shows that the implementation of differentiated instruction is still very low. The study found that the low implementation of differentiated instruction was partly due to the low self-efficacy of teachers in implementing differentiated instruction. Research shows a significant relationship between teacher self-efficacy in differentiated instruction and the implementation of differentiated instruction (Suprayogi *et al.*, 2017; DeNeve *et al.*, 2015; Dixon *et al.*, 2014), a stronger willingness to experiment with differentiated instruction (Evers *et al.*, 2002), and a higher intention to implement differentiated instruction (Wertheim & Leyser, 2002).

Although previous studies have examined teacher self-efficacy in differentiated instruction, research focusing on chemistry teachers within specific local curriculum contexts remains limited. Existing differentiated instruction self-efficacy (DISE) instruments (Bal et al., 2022; Ramli & Yusoff, 2022) were developed in different educational systems and curricular environments. They do not fully capture the pedagogical demands of the Indonesian *Kurikulum Merdeka*, particularly as implemented in local contexts such as Yogyakarta. Chemistry teachers in Yogyakarta face distinct instructional challenges, including abstract subject matter, laboratory-based learning, heterogeneous student readiness, and varying access to instructional resources. These contextual characteristics suggest the need for a context-sensitive and discipline-specific measurement instrument. To date, no validated instrument has been developed to measure chemistry teachers' self-efficacy in implementing differentiated instruction within the Yogyakarta context, creating a clear research gap addressed by this study.

Teacher Differentiated Instruction Self-Efficacy Definition and Construct

Differentiated instruction was initiated by Carol Ann Tomlinson in 1999. Her book "The differentiated classroom: responding to the needs of all learners" opened a new perspective on learning. Tomlinson criticized the One Size Fits All approach, saying that it could not meet students' individual needs. Tomlinson argued that, even though they are grouped into one class, students must be seen as individuals with different levels of readiness, interests, and learning

10.21831/pep.v29i2.91036

styles. These differences must be responded to according to need. According to Tomlinson and Strickland (2005), differentiated instruction is a teaching philosophy based on the premise that students can learn optimally when teachers accommodate student differences, including readiness levels, interests and learning profiles.

Differentiated instruction is a set of learning process options teachers design to meet students' needs (Herwina, 2021). Thus, it can be concluded that differentiated instruction is a series of decisions teachers make to determine the learning process that suits students' needs.

The concept of self-efficacy was introduced by Bandura (1977) as an assessment of one's ability to achieve a desired level of performance in a given endeavor. The concept was then contextualized in the field of education. The concept of teacher self-efficacy emerged, defined as beliefs related to the effort teachers put into teaching, their goals, their persistence when things do not go well, and their resilience in the face of unpleasant situations (Tschannen-Moran *et al.,* 1998). In the context of differentiated instruction, self-efficacy can be interpreted as a teacher's belief in the ability to analyze and interpret student data effectively in order to successfully connect or apply data findings to determine learning instructions and improve student learning (Datnow & Hubbard, 2015).

In previous studies on the construct of teacher self-efficacy, Gibson and Dembo (1984) proposed two components: personal teaching efficacy and teaching efficacy. The study by Riggs and Enochs (1990) specifically explains teacher self-efficacy in science learning, identifying two components: personal science-teaching efficacy and science-teaching outcome expectancy.

Regarding the construct of self-efficacy in the context of differentiated learning, there are previous studies that investigate the components of DISE. According to Bal *et al.* (2022), based on the DISE instrument development study, there are six components that can represent teacher self-efficacy in differentiated learning, namely the teaching process, content, learning profile, readiness, assessment, and learning environment. Ramli and Yusoff's (2022) study identified six components of DISE, namely identification of student differences, learning environment, learning content, learning process, learning products, and assessment.

Research Purposes

The purpose of this study was to develop an instrument and validate the chemistry teacher self-efficacy scale instrument in implementing differentiated instruction among chemistry teachers in Yogyakarta, Indonesia, within the context of the existing curriculum. In this study, we conducted and reported evidence of validity in the instrument measurement process. The specific research questions are as follows:

RQ1: What are the constructs of the chemistry teacher's self-efficacy for differentiated instruction instruments?

RQ2: What empirical evidence supports the reliability of the instrument?

RESEARCH METHOD

Qualitative and quantitative research approaches were used to develop and validate a new instrument to measure chemistry teachers' differentiated instruction self-efficacy within the existing curriculum in Yogyakarta, Indonesia.

Instrumentation

The framework of chemistry teacher differentiated instruction self-efficacy instrument. This study develops a new instrument to measure chemistry teacher differentiated instruction self-efficacy. This study will integrate the construct of teacher self-efficacy in general (Gibson

Copyright © 2025, Jurnal Penelitian dan Evaluasi Pendidikan, 29(2), 2025

& Dembo, 1984; Riggs & Enochs, 1990) with the construct of self-efficacy in the context of differentiated instruction (Ramli & Yusoff, 2022; Bal *et al.*, 2022). Based on the results of the study, the construction of the instrument includes six components, namely efficacy in differentiating the learning process, efficacy in differentiating learning content, efficacy in differentiating learning products, efficacy in identifying student differences, efficacy in assessment, and efficacy in differentiating the learning environment.

Item Creation

Statement items were derived from the operational definition of the construct that had been established. The study involved one professor of chemistry education in helping develop the instrument based on the components of chemistry teacher differentiated instruction self-efficacy. The researcher developed 31 initial items that covered all constructs in chemistry teacher differentiated instruction self-efficacy. All items were presented on a Likert scale with five categories: Strongly Agree (SA), Agree (A), Neutral (N), Disagree (D), and Strongly Disagree (SD). The six subscales are described as follows.

Efficacy in the differentiation of learning processes refers to the teacher's belief in his/her ability to differentiate learning processes based on students' learning styles and preferences, and the extent to which the teacher believes his/her ability to differentiate processes positively influences student performance. Efficacy in differentiating learning content refers to the teacher's belief in his/her ability to differentiate learning content based on students' learning styles and preferences, and the extent to which the teacher believes his/her belief in his/her ability to differentiate content positively influences student performance. Efficacy in differentiating learning products refers to the teacher's belief in his/her ability to differentiate learning products based on students' learning styles and preferences, and the extent to which the teacher believes his/her ability to differentiate products positively influences student performance. Efficacy in identifying student differences refers to teachers' confidence in identifying differences in students' learning profiles and learning readiness. Efficacy in assessment refers to the teachers' confidence in conducting initial assessments and evaluations to determine appropriate forms of learning. Efficacy in differentiating learning environments refers to the teacher's belief in his/her ability to tailor the learning environment to students' preferences and the extent to which the teacher believes his/her ability to differentiate learning environments positively influences student performance.

Content Validity

The construct of chemistry teacher differentiated instruction self-efficacy was built based on a strong literature review. After creating the initial instrument, the researchers recruited reviewers to review and provide feedback on the suitability of the construct to the instrument items. Two chemistry education professors conducted the review process to check the content validity of the instrument. Then, three chemistry teachers were asked to respond to the initial instrument individually to avoid their misunderstanding of the items. Their comments and suggestions were used to revise the instrument as evidence supporting the content validity of the initial instrument.

Participant

The sample of this study was 156 secondary full-time chemistry teachers (age *M*: 37.65, female: 66.67%, male: 33.33%) in Yogyakarta region who had used "*Kurikulum Merdeka*". The study visited the chemistry teachers both in person and online. The online survey was distributed to the WhatsApp groups of the chemistry subject teacher community. The sample size used in this study met the recommended criteria for factor analysis. Hair *et al.* (2019) suggest that a

10.21831/pep.v29i2.91036

minimum sample of 150–200 is generally acceptable for CFA models with moderate complexity, while Kline (2015) notes that sample sizes above 100 are sufficient when factor loadings are strong and indicators are well-specified. Classical guidelines also emphasize respondent-to-item ratios; for instance, a 5:1 ratio is considered a minimum, and a 10:1 ratio is preferred for more stable factor solutions (Costello & Osborne, 2005). In addition, Tabachnick and Fidell (2019) recommend a minimum of 200 participants, though smaller samples may be adequate when communalities are high. Importantly, MacCallum *et al.* (1999) argue that sample adequacy depends not only on absolute N but also on loading strength, communalities, and model complexity. With 156 participants for 31 indicators (ratio 5:1), supported by a high KMO value (0.919) and strong factor loadings, the sample size in this study meets established methodological benchmarks and is considered adequate for both exploratory and confirmatory factor analysis.

Data Analysis

The data were analyzed using structural equation techniques with SPSS and Amos software. We used a stepwise approach by conducting a series of exploratory factor analyses (EFA) and confirmatory factor analyses (CFA) to estimate the structure of teacher self-efficacy in implementing differentiated instruction.

Exploratory factor analysis (EFA). Exploratory factor analysis (EFA) was modelled according to Williams *et al.* (2014). Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's Test of Sphericity were used to assess the suitability of the data in conducting EFA (Hair *et al.*, 2019). KMO values of ≥ 0.60 were considered acceptable, and a significant Bartlett's Test ($p < 0.05$) indicated that the correlation matrix was appropriate for factor extraction. A series of EFA models was conducted to examine the factorial structure of each theoretically assumed dimension using the maximum likelihood extraction method, factor retention method (cumulative variance and scree plot, and selection of rotation method (orthogonal: varimax). We considered the loading factor above = 0.50 to determine the dimension map.

Confirmatory Factor Analysis (CFA). CFA aims to confirm certain relationship patterns predicted based on theory or previous analytical results (Kline, 2015; DeVellis, 2016). This study-built factors based on the EFA results by considering the construct of self-efficacy in differentiated instruction in previous literature (Gibson & Dembo, 1984; Riggs & Enochs, 1990; Ramli & Yusoff, 2022; Bal *et al.*, 2022). CFA was used to provide a statistical analysis to evaluate how well the data fit the model. In addition, the CFA analytical procedure facilitates the assessment of convergent and discriminant validity, reliability, and the measurement of individual item quality. Convergent validity was evaluated using factor loadings (≥ 0.50), Average Variance Extracted (AVE ≥ 0.50), and Composite Reliability (CR ≥ 0.70) (Hair *et al.*, 2019). Discriminant validity was assessed using the Fornell–Larcker criterion, requiring the square root of AVE for each construct to exceed its inter-construct correlations. Reliability was examined using both Composite Reliability (CR ≥ 0.70) and Cronbach's alpha (≥ 0.70). These criteria guided the psychometric evaluation of the instrument. Researchers use several types of indices to conduct goodness-of-fit index models, including absolute fit indices, relative fit indices, and parsimony fit indices (Schreiber *et al.*, 2006). Absolute fit indices included χ^2, df, RMSEA, SRMR, AGFI, and GFI. Relative fit indices included CFI, TLI, NFI, RFI, and IFI. Parsimony fit indices included PGFI, PNFI, PCFI, and χ^2/df. These indices collectively provided a comprehensive evaluation of how well the proposed model fit the empirical data. After the CFA test, we formatted the final with the constructs and their related items. Cronbach's alpha values for each subscale and the total instrument are presented as the instrument's internal consistency reliability. This CFA test uses AMOS 26.0 software.

Commented [A1]: what are the criteria?

Commented [A2R1]: Thank you for the helpful question. I have revised the Data Analysis section by adding the criteria for the Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity. Specifically, I now state that KMO values ≥ 0.60 indicate adequate sampling and that a significant Bartlett's Test (p < 0.05) confirms the suitability of the correlation matrix for factor extraction.

Commented [A3]: Criteria??

Commented [A4R3]: Thank you for the insightful comment. I have now added the explicit criteria used to assess convergent validity, discriminant validity, and reliability in the CFA section. Specifically, I included the standards for factor loadings (≥ 0.50), AVE (≥ 0.50), Composite Reliability (≥ 0.70), the Fornell–Larcker discriminant validity requirement, and Cronbach's alpha (≥ 0.70). These criteria have been incorporated into the revised manuscript to clarify the analytical procedures.

Commented [A5]: What types are there?

Commented [A6R5]: Thank you for the comment. I have revised the manuscript by specifying the types of goodness-of-fit indices used in the CFA. The revised text now clarifies that the model was evaluated using three categories of indices: absolute fit indices (χ^2, df, GFI, AGFI, RMR, RMSEA), relative/incremental fit indices (NFI, RFI, IFI, TLI, CFI), and parsimony fit indices (PGFI, PNFI, PCFI, χ^2/df).

Copyright © 2025, Jurnal Penelitian dan Evaluasi Pendidikan, 29(2), 2025
ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i2.91036

FINDINGS AND DISCUSSION

Exploratory factor analysis. The results of the Kaiser-Meyer-Olkin (KMO) test of 0.919 and Bartlett's Test of Sphericity (χ^2 = 2939.30, p = 0.000) indicate that EFA can be applied to the data. Based on this analysis, 22 items yielded six factors with eigenvalues above 1. The factor structure with an eigenvalue of one or more than one is considered stable (Tabachnick & Fidell, 2019). The scree plot results (Figure 1) show that after the sixth factor, the slope and graph form a horizontal line. Thus, it can be accepted that the scale factors consist of six dimensions.

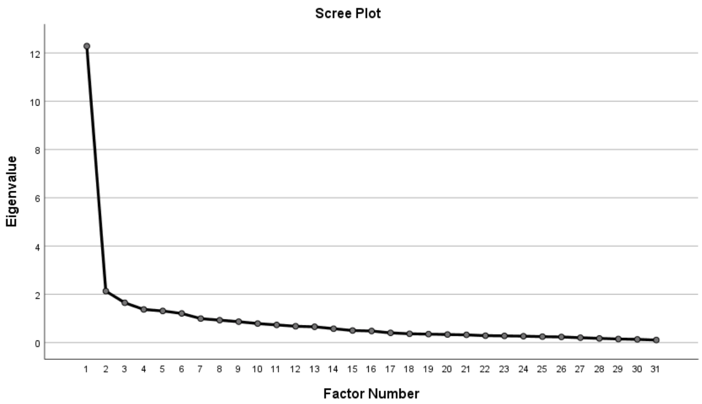

Figure 1. Scree Plot

Table 1. Exploratory Factor Analysis Result

Item Number	1	2	3	4	5	6	Mean	SD
DISE2	0.80						3.85	0.79
DISE7	0.77						3.83	0.84
DISE6	0.77						3.81	0.80
DISE4	0.71						3.83	0.76
DISE5	0.63						3.82	0.81
DISE12		0.72					3.96	0.86
DISE8		0.64					3.99	0.86
DISE11		0.56					4.03	0.79
DISE9		0.54					4.03	0.86
DISE13			0.78				3.85	0.80
DISE16			0.73				3.89	0.87
DISE15			0.67				3.89	0.84
DISE20				0.71			3.90	0.79
DISE18				0.68			3.87	0.83
DISE22				0.67			3.92	0.83
DISE19				0.65			3.92	0.79
DISE17				0.63			3.90	0.77
DISE27					0.82		3.92	0.86
DISE25					0.64		3.89	0.82
DISE23					0.54		3.88	0.75
DISE30						0.66	4.08	0.78
DISE29						0.64	4.15	0.80
Eigenvalue	12.28	1.37	1.31	2.13	1.65	1.21		
% of Variance	39.62	4.43	4.23	6.88	5.32	3.89		
Number of items	5	4	3	5	3	2		

Note. The factor loading of the items below 0.50 were not written

Copyright © 2025, Jurnal Penelitian dan Evaluasi Pendidikan, 29(2), 2025
ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i2.91036

In Table 1, the percentage of total variance explained in the scale consisting of 22 items and six factors is 64.37% of the variance, 39.62% is the sub-factor of efficacy in differentiating the learning process, 6.88% efficacy in identifying student differences, 5.32% efficacy in assessment, 4.43% efficacy in differentiating learning content, 4.23% efficacy in differentiating learning products, and 3.89% efficacy in differentiating the learning environment. The scale loading factor ranges from 0.54 to 0.82. The researchers considered omitting items with factor loadings below 0.50.

The process of naming the sub-factors of the self-efficacy scale in differentiated instruction, taking into account relevant literature. Therefore, the names of the factors are determined as "efficacy in differentiating the learning process", "efficacy in differentiating learning content", "efficacy in differentiating learning products", "efficacy in identifying student differences", "efficacy in assessment", "efficacy in differentiating the learning environment".

The first factor, "efficacy in differentiating the learning process," consists of five items. Examples of these items are "I am able to prepare several learning activities (more than one) that students can choose according to their preferences." and "My ability to differentiate the learning process based on students' abilities, learning styles, and preferences can positively affect student performance". This factor reflects teachers' confidence in adapting instructional methods, pacing, and learning activities to accommodate variations in students' readiness, interests, and learning profiles. The factor also captures the belief that adapting the learning process can positively influence students' performance. Tomlinson's framework emphasizes that process differentiation requires flexible pedagogical decision-making, which depends heavily on teacher self-efficacy. Similar factor patterns appear in Bal *et al.* (2022), although chemistry teachers in this study demonstrate stronger reliance on scaffolded instruction due to the abstract and sequential nature of chemistry concepts.

The second factor, "efficacy in differentiating learning content," consists of four items. Examples of these items are "I am able to prepare several formats of learning materials (more than one, for example video, text, PowerPoint, etc.) which can be selected according to students' learning styles" and "My ability to differentiate learning content can improve student performance". This factor represents teachers' beliefs in their ability to modify, simplify, or extend chemistry content based on student needs. It also includes teachers' confidence that differentiated content positively impacts learning outcomes. Chemistry, as a conceptually dense subject, requires strong pedagogical content knowledge (PCK), making content differentiation particularly challenging. This finding is consistent with Ramli and Yusoff (2022), who noted that content adjustment is a central predictor of DI efficacy, yet the current study reveals stronger domain-specific demands within the context of curriculum implementation among chemistry teachers in Yogyakarta.

The third factor, "efficacy in differentiating learning products," consists of three items. Examples of these items are "I am able to provide a wide selection of learning products that can be selected based on student preferences," and "My ability to provide a variety of learning product choices can improve student academic achievement". This factor captures teachers' confidence in designing varied output tasks that allow students to demonstrate understanding in multiple ways (reports, models, presentations, problem-solving tasks). Teachers also believe that offering differentiated product choices can improve academic achievement. Product differentiation aligns with constructivist principles and Bandura's emphasis on task-specific mastery beliefs. Comparable structures were reported in Gibson and Dembo (1984), but chemistry teachers in this study perceive product differentiation as a more complex task due to laboratory work, abstract reasoning, and safety considerations in chemistry classrooms.

The fourth factor, "efficacy in identifying student differences," consists of five items. Examples of these items are "I believe that every student has a different learning style" and "My ability to understand the differences in each student's characteristics is useful as capital for implementing differentiated learning". This factor indicates teachers' perceived ability to assess

10.21831/pep.v29i2.91036

and interpret student readiness, interests, prior knowledge, and learning profiles, an essential starting point for implementing differentiated instruction. It also includes recognizing that students' readiness strongly influences learning outcomes. The factor aligns with Tomlinson's (2005) assertion that differentiation begins with understanding learner variance. Consistent with Ramli and Yusoff (2022), this study highlights similar findings, though chemistry teachers show heightened attention to learning readiness due to uneven resource distribution and achievement gaps across classrooms.

The fifth factor, "efficacy in assessment," consists of three items. Examples of these items are "I need to start a new topic by assessing students' learning readiness through preliminary questions or short tests" and "I am able to assess students' achievement of differentiated instruction outcomes". This factor reflects teachers' confidence in using formative assessment strategies, interpreting assessment data, and leveraging feedback to guide differentiated instruction. This matches Bandura's view that self-efficacy heavily influences teachers' ability to engage in diagnostic and continuous assessment. The factor also aligns with previous DISE instruments, but assessment plays a more central role in this sample because ongoing evaluation is emphasized in the *Kurikulum Merdeka* implementation guidelines.

The sixth factor, "efficacy in differentiating the learning environment," consists of two items. Examples of these items are "I ensure that all educational resources are easily accessible to students by considering their needs and preferences" and "I use the same classroom layout for every learning activity". This factor represents teachers' belief in their ability to structure the physical, social, and psychological classroom environment to support differentiation, including grouping strategies, classroom management, and emotional climate. Tomlinson notes that a flexible and supportive classroom environment is foundational for effective differentiation. This factor is also consistent with patterns identified in Bal *et al.* (2022), but chemistry teachers in Yogyakarta face additional challenges related to laboratory safety, equipment availability, and diverse class sizes, making environmental differentiation a particularly salient domain.

The emergence of six distinct factors is consistent with Bandura's (1997) conceptualization of self-efficacy as a domain-specific construct, indicating that chemistry teachers' confidence in implementing differentiated instruction operates through multiple task-related components rather than as a single general belief. The strong factor loadings support the multidimensional nature of teacher self-efficacy proposed in earlier teacher efficacy research (Gibson & Dembo, 1984; Riggs & Enochs, 1990). Furthermore, the specific factors identified in this study (efficacy in differentiating content, process, and product) are aligned with Tomlinson's (2005) model of differentiated instruction, confirming that these domains are not only theoretically important but also empirically distinguishable within the context of chemistry teaching in Yogyakarta.

Based on the EFA results for the construct of self-efficacy in differentiated instruction, we tested the six-factor model using 22 items in a CFA (Figure 2). The CFA analysis was conducted using AMOS 26.0 to run the data. CFA was applied to determine whether the six-factor structure obtained from the EFA results was appropriate. The fit indices values obtained during the analysis process are presented in Table 3. The Chi-square/d*f* value (1.30) was smaller than 3. The GFI (0.88) and AGFI (0.84) indices were acceptable for absolute fit indices, and the SRMR and RMSEA indices were good. For relative fit indices, four of them (NFI, IFI, TLI, and CFI) were good enough (larger than 0.90), while the other one (RFI) was just acceptable. All three parsimony fit indices (PGFI, PNFI, and PCFI) were good (larger than 0.50). Based on the results, the values fall into the "acceptable" and "good" categories. This shows that most of the indexes show acceptable results or are in accordance with the model. In other words, it can be said that the self-efficacy scale model in differentiated instruction, consisting of 22 items, has been confirmed.

Copyright © 2025, Jurnal Penelitian dan Evaluasi Pendidikan, 29(2), 2025
ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i2.91036

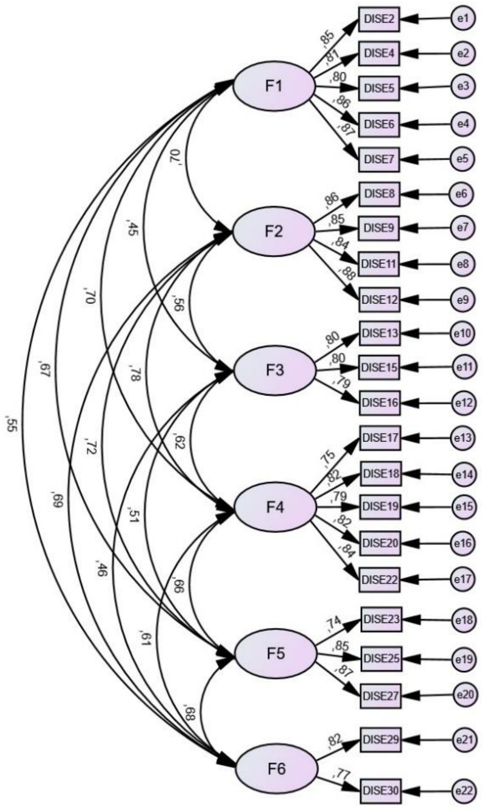

Figure 2. CFA Model

Table 3. Summary of the Goodness-of-Fit Indexes

Goodness-of-Fit Indices	Criteria	Value	Evaluation Result
Absolute fit indices			
Chi-square (χ^2)	$p > 0.05$	253.14	Poor
d*f*	-	194	-
GFI	≥ 0.90	0.88	Acceptable
AGFI	≥ 0.90	0.84	Acceptable
RMR	≤ 0.08	0.03	Good
RMSEA	≤ 0.08	0.044	Good
Relative fit indices			
NFI	≥ 0.90	0.90	Good
RFI	≥ 0.90	0.89	Acceptable
IFI	≥ 0.90	0.98	Good
TLI	≥ 0.90	0.97	Good
CFI	≥ 0.90	0.97	Good
Parsimony fit indices			
PGFI	≥ 0.50	0.67	Good
PNFI	≥ 0.50	0.76	Good
PCFI	≥ 0.50	0.82	Good
χ^2/df	≤ 3	1.30	Good

Copyright © 2025, Jurnal Penelitian dan Evaluasi Pendidikan, 29(2), 2025
ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i2.91036

Table 4 presents a summary of the CFA analysis of the model. The standardized factor loading of all items is greater than 0.45, indicating that the constructs include good convergent validity (Hair *et al*., 2019). The researchers calculated the average variance validity (AVE) values for each construct to examine if the item contributes more than the errors to the factors (Fornell & Larcker, 1981). The cut-off value of the AVEs is larger than 0.50. We found that six factors were higher than 0.50. The composite reliability (CR) values of all factors ranged from 0.77 to 0.94, exceeding the criterion value of 0.70 (Hair *et al*., 2019).

Table 4. Summary of Confirmatory Factor Analysis

Factors	Items	Standardized Factor Loading	CR	AVE
F1	DISE2	0.85	0.94	0.70
	DISE4	0.81		
	DISE5	0.80		
	DISE6	0.86		
	DISE7	0.87		
F2	DISE8	0.86	0.92	0.74
	DISE9	0.85		
	DISE11	0.84		
	DISE12	0.88		
F3	DISE13	0.80	0.84	0.63
	DISE15	0.80		
	DISE16	0.79		
F4	DISE17	0.75	0.90	0.65
	DISE18	0.82		
	DISE19	0.80		
	DISE20	0.82		
	DISE22	0.84		
F5	DISE23	0.74	0.86	0.67
	DISE25	0.85		
	DISE27	0.87		
F6	DISE29	0.82	0.77	0.63
	DISE30	0.77		

Correlations

After the CFA analysis, the researchers confirmed the instrument's final results as a six-factor construct comprising 22 items. Correlations among the six factors and the entire instrument were computed to provide further evidence of the instrument's convergent and discriminant validity. The results of Table 5 found that the correlation between each factor and the entire instrument ranged from 0.66 to 0.87. This means that the results indicate that each factor makes a significant contribution to the entire instrument. The correlations between factors ranged from 0.45 to 0.69, indicating that they served individually and related to each other. In other words, the six factors of the instrument are interdependent on each other. It is expected that these values are not negative and have a minimum value of 0.30 and above (Büyüköztürk, 2018). Thus, it can be said that the scale meets the criteria.

Reliability

Cronbach's alpha was calculated to assess the reliability of the self-efficacy scale in differentiated instruction. Table 5 shows the internal consistency reliability of the six subscales and the whole instrument. The Cronbach's alpha for F6 was 0.77, which is acceptable, while the other five factors (F1, F2, F3, F4, F5) had alphas above 0.80, indicating good reliability. The overall reliability of the instrument was 0.95, indicating that the 22-item instrument is reliable for measuring chemistry teachers' self-efficacy in implementing differentiated instruction.

10.21831/pep.v29i2.91036

The aforementioned results provide sufficient evidence to support the convergent and discriminant validity and reliability of the final instrument to measure chemistry teachers' self-efficacy in implementing differentiated instruction.

Table 5. The Reliability and Correlations Among the Total and the Six Subscales

Subscale	Item Number	Cronbach's Alpha	Correlations					
			F1	**F2**	**F3**	**F4**	**F5**	**F6**
F1	5	0.92	1					
F2	4	0.92	0.69**	1				
F3	3	0.84	0.45**	0.56**	1			
F4	4	0.90	0.69**	0.78**	0.62**	1		
F5	3	0.86	0.66**	0.72**	0.50**	0.66**	1	
F6	2	0.77	0.55**	0.68**	0.45**	0.61**	0.62**	1
Total	**16**	**0.95**	**0.83****	**0.87****	**0.66****	**0.87****	**0.78****	**0.68****

Note: ** < 0.05, ** < 0.01, *** < 0.001

This study developed and validated a new instrument to measure chemistry teachers' self-efficacy in implementing differentiated instruction. We built a valid and reliable instrument with six factors consisting of 22 items.

In the development stage, the content validity of the instrument was confirmed by carefully examining its theoretical basis (Gibson & Dembo, 1984; Riggs & Enochs, 1990; Ramli & Yusoff, 2020; Bal *et al.*, 2022). In this study, researchers involved experts in the field of chemistry education and chemistry teachers. They checked the content validity of the instrument that the researcher developed based on a strong theoretical review. Compared to previous DI self-efficacy instruments such as those developed by Ramli and Yusoff (2020) and Bal *et al.* (2022), the current instrument demonstrates a broader and more discipline-sensitive structure. For example, the inclusion of "efficacy in differentiating the learning environment" and "efficacy in assessment" reflects challenges commonly reported in chemistry instruction, such as laboratory organization and readiness-based assessment practices. These findings highlight the unique pedagogical demands faced by chemistry teachers and provide empirical evidence that subject-specific contexts shape teacher self-efficacy beliefs.

Regarding the research question (RQ1), we conducted a series of exploratory factor analyses (EFAs) and confirmatory factor analyses (CFAs) to estimate the structure of teacher self-efficacy in implementing differentiated instruction. Of the 31 items analyzed using EFA, there were several items that did not meet the criteria, resulting in 22 items with six factors. Then, we analyzed the 22 items using CFA. The analysis provided sufficient evidence at both the item and construct levels to support the overall structure of the instrument and the quality of each item and construct. The overall goodness-of-fit indices from CFA analysis can be used as evidence to support the instrument model of chemistry teacher self-efficacy in implementing differentiated instruction.

Regarding the research question (RQ2), the reliability of the instrument was based on the Cronbach's Alpha internal consistency coefficient, and both sub-factors and the entire instrument showed a value above 0.70. This indicates that the instrument is reliable (Fraenkel *et al.*, 2018).

CONCLUSION

The findings from exploratory factor analysis, confirmatory factor analysis, and correlation analysis provide sufficient empirical evidence to support the instrument's convergent and discriminant validity. The high Cronbach's coefficient alpha values show good internal

10.21831/pep.v29i2.91036

consistency and reliability of the instrument. We documented a valid and reliable chemistry teacher self-efficacy for differentiated instruction, comprising 6 constructs and 22 items, by integrating theory and data. The six-factor scale is named "efficacy in differentiating the learning process, efficacy in differentiating learning content, efficacy in differentiating learning products, efficacy in identifying student differences, efficacy in assessment, and efficacy in differentiating the learning environment". Since the instrument was validated solely within the chemistry teacher community in Yogyakarta, researchers should exercise caution when extending the findings to other regions in Indonesia or to other countries. To enhance its reliability and broader applicability, future research should focus on validating the instrument across different regions and educational settings. Additionally, the use of convenience sampling and distribution through WhatsApp groups may result in a sample that does not adequately reflect the broader population of chemistry teachers. Consequently, subsequent studies should employ more diverse or representative sampling strategies to improve population representativeness. This study contributes to the professional learning research community by developing a valid and reliable instrument to measure chemistry teacher efficacy for differentiated instruction. The instrument developed can be used as a tool to check the progress of professional development programs in implementing differentiated learning. This study also contributes to the theoretical understanding of the construct by identifying its underlying dimensions and validating these through rigorous psychometric testing using EFA and CFA. This instrument can serve as a basis or initial model. However, it is necessary to adapt and retest for different contexts, both culturally and educationally, to ensure that the instrument is relevant and effective.

Conflict of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

REFERENCES

Asriadi AM, M., Retnawati, H., & Hadi, S. (2023). Does differentiated instruction affect learning outcomes Systematic review and meta-analysis. *Journal of Pedagogical Research*, *7*(5), 18-33. https://doi.org/10.33902/JPR.202322021

Bal, A. P., Yilmaz, R., & Atas, V. (2022). Development of self-efficacy scale of differentiated instruction for teachers. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *7*(2), 93–104. https://doi.org/10.23917/jramathedu.v7i2.16204

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychol. Rev.*, *84*(2), 191–215. https://doi.org/10.1016/0146-6402(78)90002-4

Bandura, A. (1997). Self-efficacy: The exercise of control. Freeman.

Büyüköztürk, S. (2018). Sosyal bilimler icin veri analizi el kitabı. PegemA Yayınevi.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(7), 1–9. https://legacyfileshare.elsevier.com/promis_misc/teln-best-practices-in-exploratory-factor-analysis.pdf

Cruz, E. B., Schlendorf, C. P., Nicolino, P. A., & Morote, E. -S. (2019). Instructional strategies in differentiated instruction for systemic change. *Journal for Leadership and Instruction*, *18*(2), 18–22.

10.21831/pep.v29i2.91036

Datnow A., & Hubbard L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record*, *117*(4). http://dx.doi.org/10.1177/016146811511700408

DeNeve, D., Devos, G., & Tuytens, M. (2015). The importance of job resources and self-efficacy for beginning teachers' professional learning in differentiated instruction. *Teaching and Teacher Education*, *47*, 30-41. https://doi.org/10.1016/j.tate.2014.12.003

DeVellis, R. F. (2016). *Scale development: Theory and applications*. Sage Publications.

Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, *37*(2), 111-127. https://doi.org/10.1177/0162353214529042

Evers, W. J. G., Brouwers, A., & Tomic, W. (2002). Burnout and self-efficacy: A study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of Educational Psychology*, *72*(2), 227-243. https://doi.org/10.1348/000709902158865

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error, *J. Mark Res.*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Fox, J., & Hoffman, W. (2011). *The differentiated instruction book of lists*. Jossey-Bass.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2018). *How to design and evaluate research in education* (10th ed.). McGraw-Hill.

Gaitas, S., & Martins, M. A. (2017). Teacher perceived difficulty in implementing differentiated instructional strategies in primary school. *International Journal of Inclusive Education*, *21*(5), 544–556. https://doi.org/10.1080/13603116.2016.1223180

Gatune, M. R., Njagi. M. W., & Elias, M. E. (2022). Effectiveness of differentiated instruction in enhancing students' academic achievements towards learning of chemistry in secondary schools in Maara. In *D. K.* Isutsa (Ed.), *Proceedings of the 8th International Research Conference held in Chuka University from 7th to 8th Oct 2021, Chuka, Kenya, pp. 426-433*. https://repository.chuka.ac.ke/items/7dc67914-3bc4-46f7-84cd-fd1ceec817c4

George, P. S. (2005). A rationale for differentiating instruction in the regular classroom. *Theory Into Practice*, *44*(3), 185-193. https://doi.org/10.1207/s15430421tip4403_2

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*(4), 569-582. https://psycnet.apa.org/doi/10.1037/0022-0663.76.4.569

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

Heacox, D. (2012). Differentiating instruction in the regular classroom: How to reach and teach all learners. Free Spirit Publishing.

Herwina, W. (2021). Optimizing student needs and learning outcomes with differentiated learning. *Perspective of Educational Science*, 35(2). https://doi.org/10.21009/PIP.352.10

Inspectie van het Onderwijs. (2014). De staat van het onderwijs: Onderwijsverslag 2012/2013 [The state of education in The Netherlands: The 2012/2013 education report]. Inspectie van het Onderwijs.

10.21831/pep.v29i2.91036

Inspectie van het Onderwijs. (2015a). Beginnende leraren kijken terug – Onderzoek onder afgestudeerden. Part 1: De pabo [Starting teachers looking back – A study among graduates: Part 1: Teacher training]. Inspectie van het Onderwijs.

Inspectie van het Onderwijs. (2015b). De staat *van het onderwijs: Onderwijsverslag 2013/2014 [The state of education in The Netherlands: Annual report 2013/2014]*. Inspectie van het Onderwijs. Retrieved from http://www.onderwijsinspectie.nl

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Publications.

Lavania, M., & Nor, F. B. M. (2020). Barriers in differentiated instruction: A systematic review of the literature. *Journal of Critical Reviews*, *7*(6), 293–297. https://doi.org/10.31838/jcr.07.06.51

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Marlina. (2019). Panduan pelaksanaan model pembelajaran berdiferensiasi di sekolah inklusif. PLB FIP UNP.

Ramli, R., & Yusoff, N. M. (2020). Self-efficacy and differentiated instruction: A study among Malaysian school teachers. *Universal Journal of Educational Research*, *8*(4), 1252–1260. http://dx.doi.org/10.13189/ujer.2020.080416

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*(6), 625–637. https://doi.org/10.1002/sce.3730740605

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *J. Educ. Res.*, *99*(6), 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Smets, W. (2017). High quality differentiated instruction--A checklist for teacher professional development on handling differences in the general education classroom. *Universal Journal of Educational Research*, *5*(11), 2074-2080. https://doi.org/10.13189/ujer.2017.051124

Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal*, 7(7), 935-947. https://files.eric.ed.gov/fulltext/EJ854351.pdf

Suprayogi, N. M., Valcke, M., & Godwin, R. (2017). Teachers and their implementation of differentiated instruction in the classroom. *Teaching and Teacher Education*, *67*, 291-301. https://doi.org/10.1016/j.tate.2017.06.020

Tabachnick, B. G, & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson. https://www.pearsonhighered.com/assets/preface/0/1/3/4/0134790545.pdf

Tobin, R., & Tippett, C. D. (2014). Possibilities and potential barriers: Learning to plan for differentiated instruction in elementary science. *International Journal of Science and Mathematics Education*, *12*, 423–443. http://dx.doi.org/10.1007/s10763-013-9414-z

Tomlinson, C. A., & Strickland, C. A. (2005). Differentiation in practice: A resource guide for differentiating curricula, grades 9–12. ASCD.

Tomlinson, C. A., Brighton, C., Hertberg, H., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiated instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of the literature. *Journal For the Education of the Gifted* , 27(2/3), 119- 145. https://doi.org/10.1177/016235320302700203

Copyright © 2025, Jurnal Penelitian dan Evaluasi Pendidikan, 29(2), 2025
ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i2.91036

Tschannen-Moran M., Hoy A. W., & Hoy W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*, 202–248. http://dx.doi.org/10.3102/00346543068002202

van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2019). Capturing the complexity of differentiated instruction. *International Journal of Research, Policy and Practice*, 30(1), 51-67. https://doi.org/10.1080/09243453.2018.1539013

Wertheim, C., & Leyser, Y. (2002). Efficacy beliefs, background variables, and differentiated instruction of Israeli prospective teachers. *The Journal of Educational Research*, *96*(1), 54-63. http://dx.doi.org/10.1080/00220670209598791

Whitley, J., Duquette, C., Gooderham, S., Elliott, C., Orders, S., & Klan, A. (2021). Implementation of a differentiated instruction initiative: Perspectives of leaders. *Canadian Journal of Educational Administration and Policy*, 196, 49–64. https://doi.org/10.7202/1078517ar

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3), 1-13. https://doi.org/10.33151/ajp.8.3.93

Zelalem, A., Melesse, S., & Seifu, A. (2022). Teacher educators' self-efficacy and perceived practices of differentiated instruction in Ethiopian primary teacher education programs: Teacher education colleges in Amhara regional state in focus. *Cogent Education*, *9*(1), 2018909. https://doi.org/10.1080/2331186X.2021.2018909