JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)
DOI: http://dx.doi.org/10.15408/jp3i.v14i1.40028
http://journal.uinjkt.ac.id/index.php/jp3i

Measuring the Teacher's Sense of Efficacy Scale for Teachers of Children with Special Needs in Indonesia

Ratna Syifa'a Rachmahana[1,2], Muhammad Chirzin[3], Khoiruddin Bashori[4], Halim Purnomo[5]

Faculty of Psychology and Socio-Cultural Sciences, Universitas Islam Indonesia, Indonesia[1]
Faculty of Islamic Education Psychology, Muhammadiyah University of Yogyakarta, Indonesia[2]
Faculty of Ushuluddin, Religious Studies, and Islamic Thought, Sunan Kalijaga State Islamic University, Indonesia[3]
Faculty of Psychology, Ahmad Dahlan University, Indonesia[4]
Faculty of Language Studies, Muhammadiyah University of Yogyakarta, Indonesia[5]

E-mail: ratna.syifaa@uii.ac.id

Abstract

This research aims to test the validity and analysis of the Teacher's Sense of Efficacy Scale (short form) measurement tool for teachers of children with special needs in Indonesia. In this study, researchers tested three dimensions developed by Tschannen-Moran and Hoy (2001) for measuring teaching efficacy: Efficacy in Classroom Management, Efficacy in Student Engagement, and Efficacy in Instructional Strategies, with 12 items. The research subjects were 301 teachers of children with special needs. Among the subjects are 225 female and 76 male teachers, with a median age of 26-35. The sampling method utilised in this study is purposive sampling. The data analysis method used in this research is confirmatory factor analysis (CFA). The analysis results show a GFI value of 0.9 97; CFI 0.930; TLI 0.907; RNI 0.930; SRMR 0.044; and RMSEA 0.082. Further, Rasch analysis found that all 12 items of the scale were valid and used in the study. This model allows stakeholders to gain more insights into teachers' self-efficacy, especially regarding children with special needs.

Keywords: confirmatory factor analysis, Rasch analysis, teacher self-efficacy in teaching, teacher's sense of efficacy, validation.

Abstrak

Tujuan dari penelitian ini adalah untuk melakukan uji validasi dan analisis pada alat ukur Teacher's Sense of Efficacy Scale (Short Form) bagi guru-guru untuk anak berkebutuhan khusus di Indonesia. Pada studi ini, peneliti menguji tiga dimensi yang dikembangkan oleh Tschannen-Moran dan Hoy (2001) untuk mengukur efikasi mengajar, yakni Efikasi pada Manajemen Kelas, Efikasi pada Keterlibatan Siswa, dan Efikasi pada Strategi Instruksi dengan total aitem sebanyak 12. Subjek dari penelitian ini adalah 301 guru untuk anak berkebutuhan khusus, yang terdiri dari 225 guru perempuan dan 76 guru laki-laki dengan umur median 26-35 tahun. Studi ini menggunakan metode sampling purposif. Analisis data dilakukan dengan metode Confirmatory Factor Analysis (CFA) dan juga analisis Rasch. Hasil dari analisis yang telah dilakukan didapatkan nilai GFI 0.997; CFI 0.930'TLI 0.907; RNI 0.930; SRMR 0.044 dan RMSEA 0.082. Seluruh 12 aitem pada skala ini valid dan digunakan dalam studi. Model ini dapat digunakan untuk menilai efikasi diri dari para guru terutama bagi guru dengan anak berkebutuhan khusus.

***Kata kunci**: Analisis faktor konfirmatori, efikasi diri guru dalam mengajar, efikasi diri pada guru, validasi.*

Introduction

The unique characteristics of children with special needs require a deep understanding from teachers and the best methods in the teaching and learning process at school. This is because children with special needs have their special features, such as experiencing obstacles or retardation in intelligence or intellectual function, limitations in physical function, as well as social behavior and others. Therefore, students with special needs need special education services to develop their abilities optimally (Noviandari & Huda, 2018). When teachers do not understand these characteristics well, various problems can arise in accompanying children with special needs at schools (Arias et al., 2023).

Previous research conducted by Tarnoto (2016) shows various problems teachers face when accompanying children with special needs at school. The problems complained about include the lack of class support teachers, the lack of teacher competence in dealing with children with special needs, the lack of understanding of teachers about children with special needs and inclusive schools, the inappropriate educational background of teachers, and lack of teacher patience in dealing with children with special needs. Not much different from research conducted by (Faragher et al., 2020) who explained that among teachers of children with special needs in inclusive schools, the problem found was the lack of teachers trained in handling children with special needs so that this became a problem in itself for the teachers in In his class there are students with special needs. Based on these studies, it is challenging to create expectations for a conducive classroom situation, especially since limited teachers can implement education for students with special needs.

Researchers also conducted an initial survey via Google Forms on several teachers accompanying students with special needs (*Anak Berkebutuhan Khusus,* hereinafter referred to as ABK) in Yogyakarta to discover the problems teachers face in accompanying ABK. Teachers were asked to answer the question, namely, "What obstacles did you experience as a special education teacher"? Based on the answers given, the researcher can conclude that teachers reveal that they do not yet have effective strategies and appropriate learning methods for ABK students with various characteristics. The lack of special assistant teachers is also an obstacle to providing adequate support. Teachers who act as subject and special assistant teachers face significant challenges, especially regarding class arrangements between ABK students and other students in inclusion classes.

Overall, the problems that have been described highlight the challenges faced by teachers in accompanying ABK students. Regulation of the Minister for State Apparatus Empowerment and Bureaucratic Reform (Permen PAN-RB) No. 16 of 2009, concerning "Teaching Profession and Credit Scores", Article 3 states that Types of Teachers based on their characteristics, duties and activities include: i) Class Teachers, ii) Subject Teachers, iii) Guidance and Counseling Teachers/Counselors. These three types of teachers can be placed in public and special schools (*Sekolah Luar Biasa,* hereinafter referred to as SLB). Meanwhile, Article 13 paragraph (4), which regulates the "additional duties" of the three types of teachers, mentions one of the additional duties of teachers, namely, becoming a special supervisor in an educational unit that provides inclusive education. This means that teachers in inclusive education can be teachers who do not have to have adequate competence (Handayani & Rahadian, 2013).

Starting from this, teachers need self-efficacy to increase the success of teachers accompanying children with special needs. Self-efficacy refers to an individual's belief in their abilities and an assessment of the actions needed to achieve their level of performance (Fabelico & Afalla, 2020; Skaalvik & Skaalvik, 2007). Teacher self-efficacy in teaching or teaching self-efficacy is defined as a teacher's belief or assessment of their abilities (Tschannen-Moran & Hoy, 2001). According to Bandura (1977), teaching self-efficacy correlates with their class performance, motivation for self-improvement, and attitudes towards teaching. An insight into this subject can be perceived through TSES adaptation, which will encourage relevant stakeholders to determine what training is necessary for teachers. Teachers' self-efficacy will strengthen their ability to deal with special needs students.

Confidence in the teacher will determine how much the teacher's ability to accept challenges influences the actions the teacher will take in class, the efforts the teacher will make to face challenges

http://journal.uinjkt.ac.id/index.php/jp3i

and how long the teacher will be able to persevere in facing these challenges (Putri & Fakhruddiana, 2018). Teachers with high self-efficacy can better implement a quality learning environment by creating learning plans to improve student abilities, involving students and managing disruptive student behaviour in class fully (Zee & Koomen, 2016). Teachers with high teaching self-efficacy have also been proven to contribute to improving teachers' psychological well-being, job satisfaction, commitment to work and reducing levels of stress and burnout (Aloe et al., 2014; Collie et al., 2012; Klassen & Chiu, 2011).

Research related to teaching efficacy is increasingly developing in various countries. The instruments that measure teacher efficacy are: the Teachers' Sense of Efficacy Scale (TSES), which was developed by Tschannen-Moran & Hoy (2001). Measuring teachers' efficacy was first initiated by Bandura (1977) by introducing four indicators of teacher efficacy: achievement of learning objectives, direct experience, verbal encouragement and emotional stimulation. Then, Tschannen-Moran and Hoy (2001) tested the instrument. This led to proposing the Ohio State Teacher Efficacy Scale (OSTES) instrument model, which is currently called the Teachers' Sense of Teacher Efficacy Scale (TSES). The TSES instrument has two formats: a short format comprising 12 items and a long format comprising 24 items. Each item results from a description of three dimensions: student engagement, instructional strategies and classroom management. The student engagement dimension describes the teacher's task in increasing student involvement in school activities. Furthermore, the instructional strategies dimension describes the teacher's task in implementing effective teaching strategies. The classroom management dimension describes the teacher's duties in managing the classroom.

Testing of the TSES measuring tool continues to develop to measure teaching efficacy occasionally (Kleinsasser, 2014). The TSES measuring tool has been tested with teacher research respondents in various countries (Klassen & Chiu, 2011; Maulana et al., 2020). However, most of these instruments were developed based on suitability for Western contexts. There is a lack of published research studies on teacher teaching efficacy in Indonesia. As far as the research has been carried out, researchers have not found the use or testing of the TSES measuring instrument specifically in the population of ABK teachers in Indonesia. Existing research involves elementary school teachers who live and teach in the DKI Jakarta area using the Item Response Theory approach (Maulana et al., 2020). Apart from this, it strengthened researchers to adapt the TSES measuring tool that Tschannen-Moran & Hoy (2001) developed for ABK teachers in Indonesia.

Adaptation of measuring instruments is research carried out to make original measuring instruments originating from abroad (in foreign languages) so they can be used in Indonesia. This adaptation process not only translates the language so that it can be read and understood by respondents, but is also reliable and valid according to the purpose of the measuring instrument (Azwar, 2019). In validating TSES, previous studies used several advanced psychometric methods such as confirmatory factor analysis (e.g., Gálvez-Nieto et al., 2023) or exploratory factor analysis (e.g., Kang et al., 2019; Barendse, 2005). However, another study stated that implementing modern psychometric methods will provide more detailed information about the psychometric properties of an instrument (Rahayu et al., 2024).

One of the prominent methods is the Rasch model (Wright, 1996). Applying the Rasch model combined with CFA, called the "combined approach," can be viewed from two opposing perspectives. First, research states that combining the two methods will complement each other for detailed instrument analysis (Chachamovich et al., 2008; Suseno et al., 2022). Second, research stating that CFA and Rasch come from two different philosophical views of measurement and cannot be combined in analysing instruments.

However, in the present study, we do not enter the debate and follow the approach of combining CFA and the Rasch measurement model to analyse the Indonesian TSES. Therefore, this study aims to evaluate the psychometric properties of the Indonesian TSES instrument using the CFA and Rasch model methods on a sample of teachers of special needs children. The novelty of this study means that it also adds to the first Rasch analysis contribution to the research on TSES.

Methods

Research Subject

The research aims to test the validation and analysis of the TSES for teachers of children with special needs in Indonesia. This research involved 301 teachers of children with special needs who were selected using the *accidental sampling method*. Most of the teachers of children with special needs in this study were 71 (23.6%), teachers of children with special needs in special schools and 154 (51.2%) at the elementary school level. Furthermore, most of the teachers of children with special needs in this study were women, numbering 225 (74.8%); Non-Civil Servants 264 (87.7%); the age group 26-35 years 153 (50.4%), and undergraduates (S1) numbered 289 (96%). Detailed demographics of research subjects can be seen in Table 1.

Table 1. Demographics of research subjects

Demographic Variables	Frequency (n)	Percentage (%)
Gender		
Man	76	25.249
Woman	225	74.751
Age Group		
< 25 years	53	17.608
26-35 years old	153	50.381
36-45 years old	48	15.947
45 years old	47	15.615
Last education		
High School (SMA)	5	1.661
Diploma	6	1.993
Bachelor degree)	289	96.013
Master (S2)	1	0.332
Teacher Status		
Civil servants	37	12.292
Non-Civil Servants	264	87.709
School Level		
Elementary School (SD)	154	51.163
Junior High School (SMP)	125	41.528
High School (SMA)	22	7.309
Type of School		
Inclusive Schools	221	73.422
Special School (SLB)	62	20.598
Others	18	5.980

Sources: Personal Data (2025)

Research Instruments

Researchers adapted the TSES developed by Tschannen-Moran & Hoy (2001). *The Teacher's Sense of Efficacy Scale* consists of two formats: *short form* with 12 items and *long form with* 24 items. In this research, researchers adapted the TSES shown in Table 2. This scale consists of 3 dimensions and 12 items, namely *Efficacy in Classroom Management* (4 items), *Efficacy in Student Engagement* (4 items), and *Efficacy in Instructional Strategies* (4 items). This measuring instrument is a Likert scale with five answer choices, namely 5 (very capable) to 1 (very unable). An example of a statement for this measuring tool is "How able are you to calm down students who are disruptive or noisy?".

Table 2. Blueprint *Teacher's Sense of Efficacy Scale*

Dimensions	Items
Efficacy in Classroom Management	1, 6, 7, 8
Efficacy in Student Engagement	2, 3, 4, 11
Efficacy in Instructional Strategies	5, 9, 10, 12
Number of Items	12

Sources: Personal Data (2025)

Table 3. Items for the adapted version of *Teacher's Sense of Efficacy Scale*

Number	Item	Dimension	Polarity
1.	Seberapa mampu Anda mengendalikan perilaku mengganggu yang dilakukan siswa di kelas	Efficacy in Classroom Management	Favorable
2.	Seberapa mampu Anda memotivasi siswa yang menunjukkan minat rendah pada pekerjaan sekolah?	Efficacy in Student Engagement	Favorable
3.	Seberapa mampu Anda membuat siswa percaya bahwa mereka dapat mengerjakan tugas sekolah dengan baik?	Efficacy in Student Engagement	Favorable
4.	Seberapa mampu Anda membantu siswa untuk menghargai proses pembelajaran?	Efficacy in Student Engagement	Favorable
5.	Sejauh mana Anda dapat menyusun pertanyaan yang baik untuk siswa Anda?	Efficacy in Instructional Strategies	Favorable
6.	Seberapa mampu Anda membuat siswa mengikuti aturan kelas?	Efficacy in Classroom Management	Favorable
7.	Seberapa mampu Anda menenangkan siswa yang mengganggu atau berisik?	Efficacy in Classroom Management	Favorable
8.	Seberapa mampu Anda membangun sistem manajemen kelas dengan setiap kelompok siswa?	Efficacy in Classroom Management	Favorable
9.	Seberapa mampu Anda menggunakan strategi penilaian yang bervariasi?	Efficacy in Instructional Strategies	Favorable
10.	Sejauh mana Anda dapat memberikan penjelasan atau contoh alternatif ketika siswa bingung?	Efficacy in Instructional Strategies	Favorable
11.	Seberapa mampu Anda membantu keluarga agar anak-anak mereka dapat berprestasi di sekolah?	Efficacy in Student Engagement	Favorable
12.	Seberapa mampu Anda dalam menerapkan alternatif strategi di kelas Anda?	Efficacy in Instructional Strategies	Favorable

Sources: Personal Data (2025)

Research Procedure

The stages of adapting *the Teacher's Sense of Efficacy Scale* are as follows: in the first stage, the

measuring instrument is translated into Indonesian, and then the translation results are discussed again to determine the synthesis results. The next stage, the synthesis results are translated back into English. Experts then discuss all translation results to analyse the equality between English and Indonesian items. The researcher did the translation under the supervision of the promotor and experts in psychology. The items that have been translated can then be tested on research participants, namely teachers in Special Schools (SLB) and inclusion schools, by directly distributing research questionnaires (Azwar, 2019).

The researcher first briefly introduces the research subjects, then the researcher's identity, research objectives, research procedures, and confidentiality are explained, and there is a written consent *section*. Research subject participation is entirely voluntary and anonymous. Research subjects willing to participate fill in the research measuring instruments. On average, research subjects took 10-15 minutes to fill in the research measuring instrument. After collecting the data, the researcher conducted a reliability test by calculating the McDonald's Tau coefficient with a minimum limit of 0.7 (Viladrich et al., 2017). The analyses were conducted to ensure the scale's internal consistency would pass the required threshold (Azwar, 2019).

The reliability test results obtained a result of 0.902 (>0.70), with the result being *that the Teacher's Sense of Efficacy Scale is* considered suitable for measuring the efficacy of teachers of children with special needs. This result indicates that the measuring instrument is considered significant (accepted) and can be used as a research instrument (Viladrich et al., 2017).

Data Analysis

Construct validity using CFA

This research employed JASP version 0.19 software for descriptive statistics and Confirmatory Factor Analysis. The estimation method of this study is Maximum Likelihood Estimation (MLE), mainly used for testing models with specific distributional assumptions. This study conducted a Confirmatory Factor Analysis (CFA) analysis for validation testing. (Umar & Nisa, 2020) explain that CFA can be used to test (confirm) the extent to which all items of a measuring instrument can measure or provide information related to what is to be measured. In the CFA analysis the model is declared fit if it meets several requirements, including (a) the *RMSEA value* must be <0.08; (b) the model is considered adequate if it has a goodness of fit index, namely *a CFI/TLI/RNI value* of 0.90 or more; (c) the average value of the standardised residual between the observed and predicted covariance, namely the *SRMR value* must be <0.10; (d) the minimum factor loading value is 0.25 (Hu & Bentler, 1998). In this study, the step-by-step analysis flow for conducting CFA was followed from the steps taken by Putra and El Fahmi (2024), which follows JARS quantitative (Appelbaum et al., 2018).

Further, Multigroup CFA was also performed to test whether gender-invariance (1 = female; 0 = male) was achieved. However, only configural invariance tested in this study indicates that the 3-factor model holds and fits the male and female respondents (Luong & Flake, 2023).

Rasch Analysis

After the CFA was conducted, Rasch analysis was performed using Jamovi version 2.3.28.0 for the Rasch analysis. However, because Jamovi did not provide Principal Component Analysis of Residuals (PCAR; Wright, 1996), Winsteps was also used to provide the PCAR testing with the criteria of >40% as an indication of the fulfilment of unidimensionality assumptions (Holster & Lake, 2016). This study's estimation method of Rasch analysis is the Unconditional Maximum Likelihood Estimation (UCON), which is mainly used for testing the Rasch models as fixed effect models (Hayat et al., 2020). In this study, Rasch analysis was carried out using a step-by-step procedure that was reported on (Rahayu et al., 2024). Given that the instrument has the same number of response categories, the Rating Scale Model (RSM; Andrich, 1978) was used.

In this study, RSM were performed using three separate analyses due to the multidimensional nature of the scale (i.e., 3-factor model). Rahayu et al. (2024) stated that this approach can be used and

has been previously used in some studies (Pichardo et al., 2018) if the researcher has limitations regarding implementing the multidimensional Rasch model. However, it should be noted that the estimation would be more accurate if the multidimensional Rasch model were used.

Further, after the PCAR was performed for each aspect, local independence testing was performed using Q3 statistics with the criteria of <0.30 indicating the fulfilment of the local independence assumption (Christensen et al., 2025). Last but not least, the item fit for individual items was also reported through Infit MNSQ and Outfit MNSQ, where the values in the range of 0.6-1.4 indicate that the items fit the model (Bond, 2015).

Results

Confirmatory Factor Analysis (CFA)

Before the CFA, an assumption analysis of univariate normality was conducted to determine whether the data collected could be treated as continuous. The descriptive statistics results can be seen in Table 4.

Table 4. Descriptive statistics for all items

	Mean	SD	Skewness	SE of Skewness
E1	4.013	0.432	-0.176	0.140
E2	3.983	0.486	-0.218	0.140
E3	4.086	0.431	0.224	0.140
E4	4.113	0.490	0.089	0.140
E5	4.063	0.509	0.106	0.140
E6	4.070	0.515	-0.043	0.140
E7	4.063	0.509	0.106	0.140
E8	3.917	0.526	-0.096	0.140
E9	3.890	0.540	-0.339	0.140
E10	4.066	0.478	-0.729	0.140
E11	3.821	0.567	-0.340	0.140
E12	3.954	0.481	-0.130	0.140

Sources: Personal Data (2025)

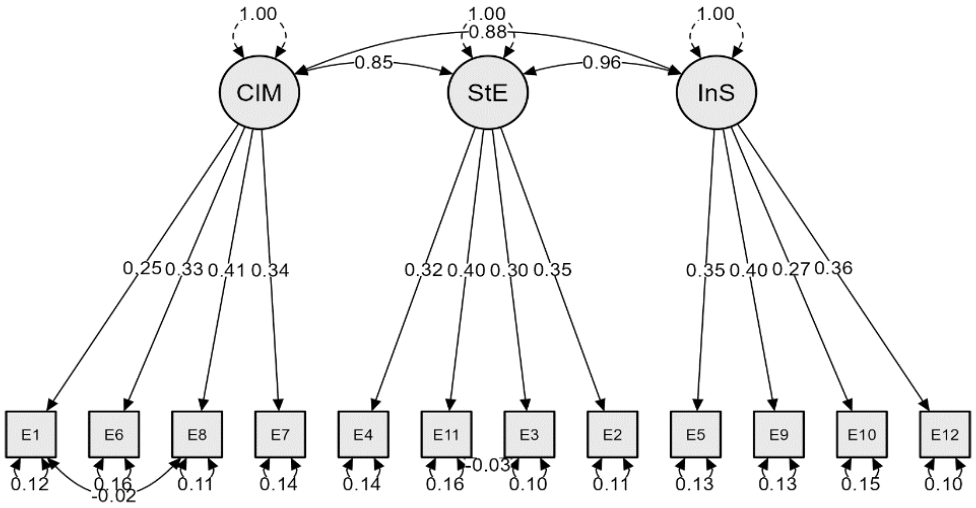
As shown in the table above, all items have skewness values ranging from -1 to +1. This finding indicates that the normality assumption supporting CFA application is met. Therefore, CFA treats all items as continuous variables, i.e. Suseno et al (2022). The SE of skewness is all uniform (SE = 0.140), in line with an analysis phase of continuous data.

After the univariate normality for all items was confirmed, the goodness-of-fit testing was performed in the first phase of CFA (see Table 5). Based on CFA, it was found that the GFI value is 0.997; CFI 0.930; TLI 0.907; RNI 0.930; SRMR 0.044; and RMSEA 0.082. Based on these statistics and indices, it can be concluded that the model has an acceptable fit.

Table 5. Fit Indices Teacher's Sense of Efficacy Scale

Goodness of Fit Index	Value
Goodness of fit index (GFI)	0.997
Comparative Fit Index (CFI)	0.930
Tucker-Lewis Index (TLI)	0.907
Relative Noncentrality Index (RNI)	0.930
Standardised root mean square residual (SRMR)	0.044
Root mean square error of approximation (RMSEA)	0.082

Sources: Personal Data (2025)



Sources: Personal Data (2025)

Figure 1. *Teacher's Sense of Efficacy* Multidimensional Model

The latent correlation between dimensions in Table 6 shows that the *Classroom Management dimension* with *Student Engagement* has a correlation estimate value of 0.0857 (p <0.001). T*he Classroom Management* dimension with *Instructional Strategies* has a correlation with the Estimate value 0.876 (P<0.001), and the *Student Engagement* dimension with *Instructional Strategies* correlates with an Estimate value of 0.989 (P<0.001), meaning dimensions *Classroom Management*, *Student Engagement,* and *Instructional Strategies* can is said to be significant and can stand alone.

Table 6. Factor Covariance Teacher's Sense of Efficacy Scale

Dimensions	Estimate	Std.Error	Z-value	P
Classroom Management↔ Student Engagement	0.857	0.035	25.562	< .001
Classroom Management↔ Instructional Strategies	0.876	0.033	27.138	< .001
Student Engagement↔ Instructional Strategies	0.989	0.027	36.043	< .001

Sources: Personal Data (2025)

Table 7 shows that all items in *the Classroom Management dimension* have estimated values with a range of 0.254-0.415, and all items in the *Student Engagement dimension* have estimated values with a range of 0.298–0.388. Furthermore, all items (E5, E9, E10, and E12) in the *Instructional Strategies dimension* have estimated values in the range 0.272–0.398. Factor loadings of each item toward the presupposed

dimension have fulfilled the required threshold of 0.250. All items can be used as measurements and adequately represent their respective factors.

Table 7. Factor Loadings

Items	Dimensions	Estimate	Std. Error	Z-value	P
E1	*Classroom Management*	0.254	0.025	9.716	<0.01
E 6	*Classroom Management*	0.329	0.029	11.559	<0.01
E 8	*Classroom Management*	0.415	0.028	11.989	<0.01
E 7	*Classroom Management*	0.360	0.028	14.560	<0.01
E 4	Student Engagement	0.319	0.026	13.334	<0.01
E 11	Student Engagement	0.388	0.024	11.836	<0.01
E 3	Student Engagement	0.298	0.026	12.065	<0.01
E 2	Student Engagement	0.344	0.030	12.885	<0.01
E 5	*Instructional Strategies*	0.353	0.027	13.239	<0.01
E 9	*Instructional Strategies*	0.398	0.028	14.068	<0.01
E 10	*Instructional Strategies*	0.272	0.027	10.227	<0.01
E 12	*Instructional Strategies*	0.359	0.025	14.281	<0.01

Sources: Personal Data (2025)

With the statistical evidence that the 3-factor model fits the data (both at the factor level and item level), further analysis was conducted by including gender as a covariate and testing the MG-CFA model. Only the configural invariance model was tested and reported in the following subsection.

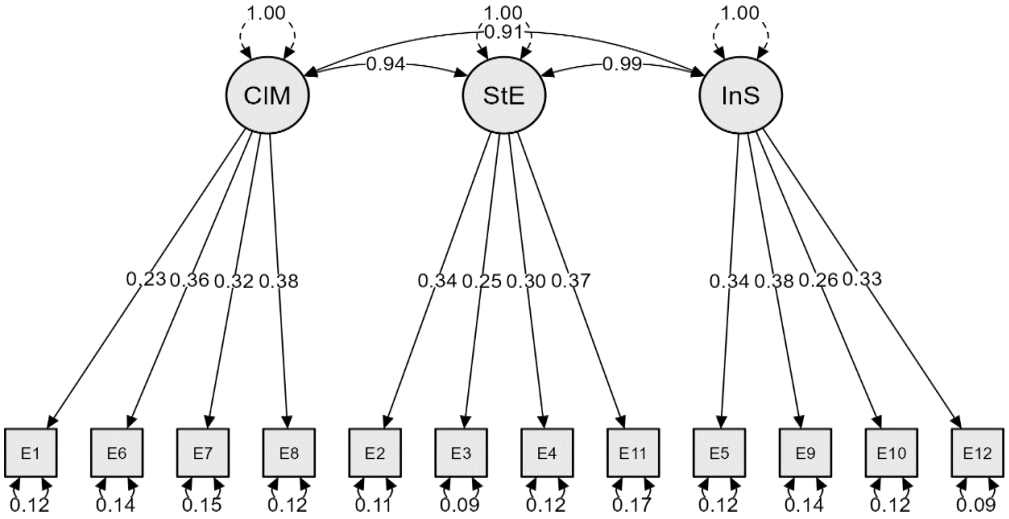
Configural Invariance Testing Using MG-CFA

To enrich the understanding of the scale, we conducted a Multigroup Confirmatory Factor Analysis by gender with configural invariance testing. The model, however, achieved poored model fit indices (GFI: 0.995, CFI: 0.880, TLI: 0.845, RNI: 0.880, SRMR: 0.057, RMSEA: 0.111) presumably since unequal number of responses between groups with female participants disproportionately outnumber male participants by 225 to 76 respectively (see Table 8).

Table 8. Fit Indices Teacher's Sense of Efficacy Scale Multigroup Confirmatory Factor Analysis by Gender

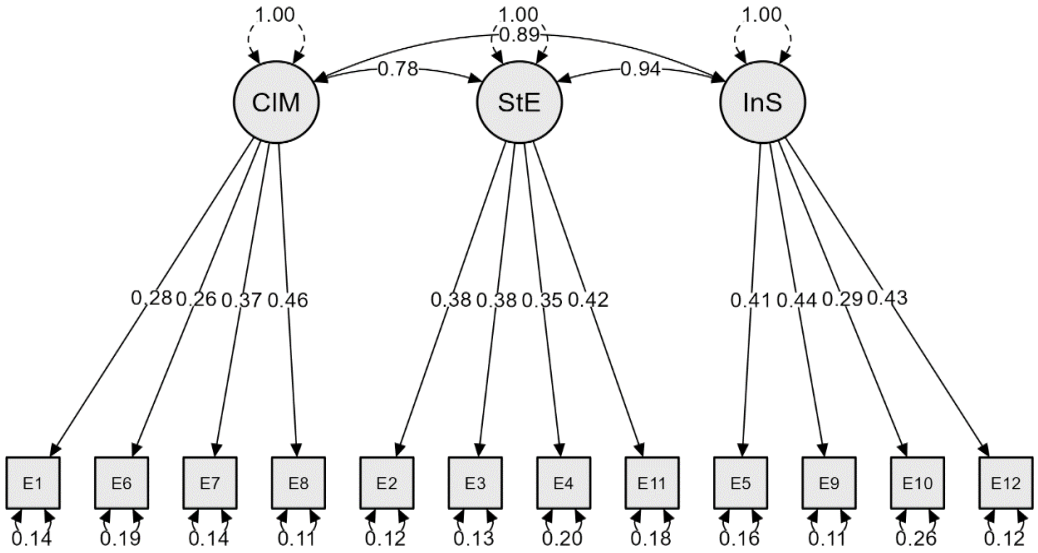
Goodness of Fit Index	Value
Goodness of fit index (GFI)	0.995
Comparative Fit Index (CFI)	0.880
Tucker Lewis Index (TLI)	0.845
Relative Noncentrality Index (RNI)	0.880
Standardised root mean square residual (SRMR)	0.057
Root mean square error of approximation (RMSEA)	0.111

Sources: Personal Data (2025)



Sources: Personal Data (2025)

Figure 2. *Teachers' Sense of Efficacy* Model in Female Participants



Sources: Personal Data (2025)

Figure 3. *Teachers' Sense of Efficacy* Model in Male Participants

This finding suggests that measurement invariance does not hold across genders. On the other hand, configural invariance is the lowest standard of measurement invariance testing. Therefore, we did not continue the MG-CFA analysis to other levels (e.g., residual invariance) because it would not be met. Ultimately, we continued the analysis with RSM to provide a detailed item-level analysis with a different measurement invariance perspective from CFA (Rahayu et al., 2024)

Rasch Rating Scale Analysis

In the first phase, unidimensionality is confirmed based on the criteria of the raw variance explained by measures, which should be greater than 40% (Holster & Lake, 2016) for each aspect. The results of the PCAR of the *classroom management* items showed that measures explained 40.2% of the raw variance. For the *student engagement* items, 42.6% of the raw variance was explained by measures. Additionally, for the *instructional strategy* items, 43.8% of the raw variance was explained by measures. Based on this finding, the unidimensionality assumptions for each aspect were achieved.

In the second phase, the Q3 Correlation Matrix of TSES in Table 9 showed no values above 0.300; therefore, item responses were independent after accounting for the latent trait. This finding confirmed the fulfilment of the local independence assumptions of the scale.

Table 9. Q3 Correlation Matrix of TSES

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11
E1	-										
E2	-0.044	-									
E3	-0.143	0.206	-								
E4	-0.252	-0.166	0.075	-							
E5	-0.109	0.061	-0.122	-0.090	-						
E6	0.003	-0.113	-0.056	-0.060	0.025	-					
E7	0.146	-0.150	-0.184	-0.217	-0.018	0.009	-				
E8	-0.119	-0.069	-0.140	-0.055	-0.069	-0.024	-0.045	-			
E9	-0.184	-0.073	-0.162	-0.024	-0.072	-0.238	-0.122	-0.009	-		
E10	-0.193	-0.199	-0.014	0.129	-0.141	-0.194	-0.145	-0.112	-0.160	-	
E11	-0.119	-0.115	-0.183	-0.095	-0.113	-0.192	-0.179	-0.047	0.108	0.038	-
E12	-0.205	-0.047	-0.163	-0.100	-0.078	-0.192	-0.130	0.020	0.227	-0,069	0.034

Sources: Personal Data (2025)

In the third phase, at the item-level, the results for the general scale found that the data fit the Rasch RSM ($p<0.001$). In Table 10, all items were considered productive for measurement because the infit and outfit mean squares were between 0.675 and 1.193, where the item fit statistics were within the range of 0.6 – 1.4 (Bond, 2015), which implied that all items fit the RSM. This finding confirmed the construct validity of the TSES using a three-separate analysis approach.

Table 10. Item statistics of the Polytomous Rasch model

	Aspects	**Measure**	**S.E.Measure**	**Infit**	**Outfit**
E1	Efficacy in Classroom Management	-4.37	0.155	1.001	0.952
E2	Efficacy in Student Engagement	-4.16	0.153	0.875	0.782
E3	Efficacy in Student Engagement	-4.91	0.159	0.820	0.708
E4	Efficacy in Student Engagement	-5.11	0.159	1.057	0.997
E5	Efficacy in Instructional Strategies	-4.74	0.158	0.988	0.955
E6	Efficacy in Classroom Management	-4.79	0.158	1.193	1.167

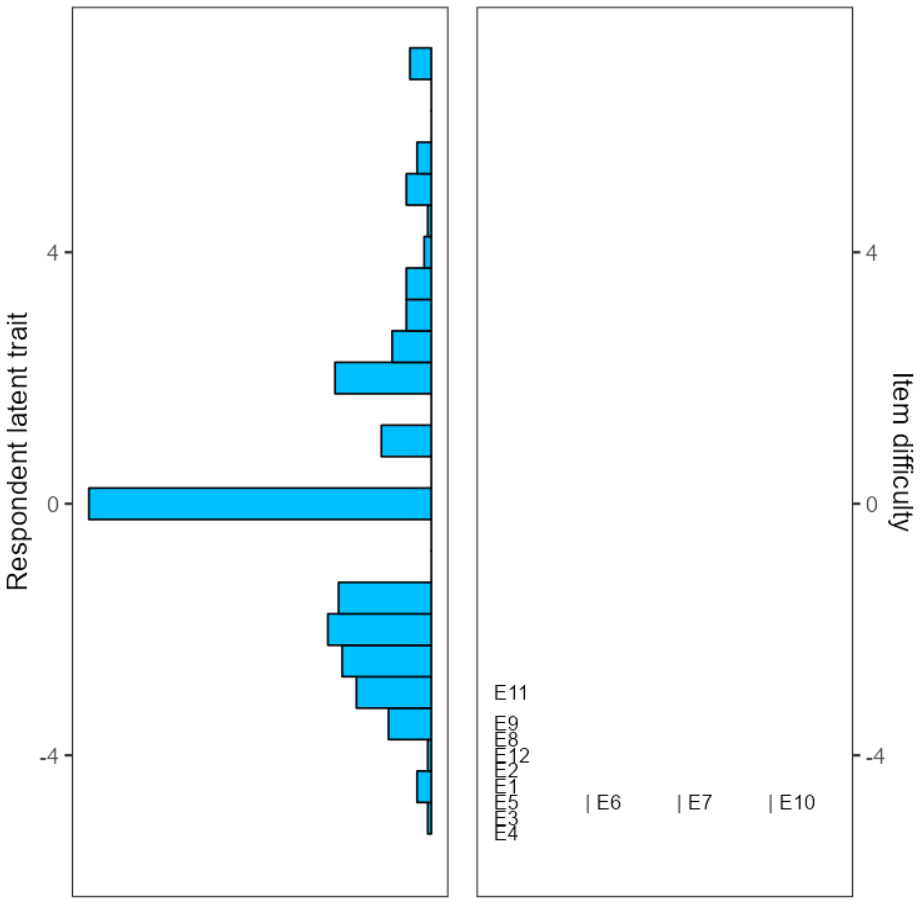
	Aspects	Measure	S.E.Measure	Infit	Outfit
E7	Efficacy in Classroom Management	-4.74	0.158	1.156	1.167
E8	Efficacy in Classroom Management	-3.71	0.147	0.882	0.841
E9	Efficacy in Instructional Strategies	-3.53	0.145	0.935	0.869
E10	Efficacy in Instructional Strategies	-4.76	0.158	1.137	1.077
E11	Efficacy in Student Engagement	-3.11	0.139	1.005	1.046
E12	Efficacy in Instructional Strategies	-3.95	0.150	0.783	0.675

Note. Infit= Information-weighted mean square statistic; Outfit= Outlier-sensitive mean square statistic.

Sources: Personal Data (2025)

Furthermore, on the Wright Map in Figure 2, it can be inferred that the respondent's latent trait was relatively widespread, more than the item distribution. The distribution of the item only covered the limited range of the person measure in the lowest continuum. These findings indicate that the items were too easy for the respondents. However, these findings do not affect the validity of all TSES items.

http://journal.uinjkt.ac.id/index.php/jp3i



Sources: Personal Data (2025)

Figure 4. Wright Map for Rasch Analysis of TSES

Discussion

The results showed that *the Teacher's Sense of Efficacy Scale,* developed by Tschannen-Moran and Hoy (2001), shows a model with three factors/dimensions to measure teacher self-efficacy in teaching children with special needs in Indonesia. Almost every indicator shows goodness-of-fit in the model, except for RMSEA, a crucial indicator for goodness-of-fit. This could be due to the smaller sample pool, which should be considered for future studies to include a larger sample pool. However, according to (MacCallum et al., 1996), a value of RMSEA between 0.080 and 0.100 can indicate a moderate fit in the model. However, further cautions should be applied before using the scale. A larger sample pool may provide a better Confidence Interval (CI).

Three-factor models derived from the study by Tschannen-Moran & Hoy (2001) demonstrate adequate internal consistency and are interconnected. The three dimensions are *the Classroom Management dimension* measured by four items (E1, E6, E7, and E8); *the Student Engagement* dimension is measured with four items (E2, E3, E4, and E11); and *the Instructional Strategies dimension* is measured with four items (E5, E9, E10, and E12). *The Teacher's Sense of Efficacy Scale* reached a reliability level of 0.902. It is proven reliable because it exceeds the minimum McDonald's Tau value *of ≥ 0.7.* Viladrich et al. (2017) explain that the minimum value of McDonald's Tau or Omega coefficient in the reliability test is 0.7. These results can be interpreted as an insight into Teaching Self-efficacy. Between the items in the study, there can then be prioritisation of which part of teaching self-efficacy can be regarded as more important to the stakeholders to develop.

We also conducted Rasch analysis to ensure the scale's robustness and how it may gauge the credibility of both the respondents and items. All items were productive for measurement, with infit and

outfit mean squares ranging between 0.675 and 1.193—well within the acceptable range of 0.6 to 1.4 (Bond & Fox, 2015), indicating that no items were misfitting, overly predictable, or exhibiting erratic responses. Additionally, the Q3 correlation matrix revealed no residual dependencies above 0.300 (see Christensen et al., 2017), supporting the assumption of local independence after accounting for the latent trait. One-dimensionality was confirmed through the unexplained variance in the first contrast (8.4%, below the 20% threshold), ruling out the presence of secondary constructions. Finally, the Wright Map illustrated widespread respondent abilities, suggesting a diverse sample. At the same time, the narrow distribution of item difficulties implied limited discrimination across responses—a pattern attributable to the Likert scale's design, which prioritises attitudinal consistency over extreme variability in difficulty.

The Teacher's Sense of Efficacy Scale can be used to measure teacher self-efficacy in teaching the population of teachers of children with special needs in Indonesia. Other studies using the same scale also showed the robustness of the scale and its model (e.g., Galvez-Nieto et al, 2023; Antoniou et al, 2023). In Galvez Nieto et al (2023), the model with bifactorial exploratory structural equation modelling demonstrated the best fit, while Antoniou et al (2023) used the same scale to do a correlational study with occupational stress, with a high overall Cronbach's alpha score (0.940). However, there are still several limitations to adapting this measuring instrument. The limitations of this research are: The distribution of data is uneven because the research subjects are special needs education teachers in special needs and inclusion schools at various levels (elementary, middle and high school). Even data distribution is needed for generalisation because teachers of children with special needs have different work pressures at each level, and more information is needed for future researchers. It is hoped that a larger research sample will be conducted from a wider coverage area in Indonesia. This is an effort to increase the generalisation of teacher self-efficacy in teaching children with special needs in Indonesia.

The multigroup confirmatory factor analysis, according to gender, was conducted using a lesser fit model with fewer fit criteria. This might be because the unequal number of responses in groups with female participants heavily outnumber male participants. It is pertinent to include a proportionate sample if there would likely be a better model fit; however, it is also important to note the disproportionate gender representation in the teaching profession, especially teaching children with special needs. In future studies, more participants might mitigate this error and showcase better model fit. The correlation between dimensions differs in both groups, with a notable difference in the relationship between *Class Management* and *Student Engagement,* where the result is significantly higher in the female group (0.94) than in their male counterpart (0.78). Previous research (e.g. Rani & Jain, 2023; Sarfo et al., 2015) suggests no significant difference between males and females in teacher efficacy; therefore, future studies must delve deeper into this matter.

Implications for teaching and learning

This study, which provides empirical validity evidence of the Indonesian TSES, has implications for teaching and learning that align with the study conducted by Hayat (2024). The main implication is that this instrument can be used as a diagnostic tool to determine the level of self-efficacy of teachers of children with special needs. The diagnostic tool aims to determine whether the level of teacher efficacy is optimal. The next implication is that this instrument can be used as a reference in teaching psychometrics regarding the combined approach, which combines CFA and Rasch models (Chachamovich et al., 2008). As discussed in previous studies, the combined approach is rarely used in Indonesia. Therefore, teaching about the combined approach can use the present study article as an example.

Conclusion

The research aims to test the validation and analysis of the Teacher's Sense of Efficacy Scale measuring tool for teachers of children with special needs in Indonesia, so that it can be used to measure teachers' self-efficacy abilities in teaching children with special needs in Indonesia. The instrument used in this study is multidimensional with three distinct constructs according to the factor loadings. Based on

the model analysis results using *Confirmatory Factor Analysis* (CFA), a GFI value of 0.9 97; CFI 0.930; TLI 0.907; RNI 0.930; SRMR 0.044; and RMSEA 0.082. This means that despite the Teacher's *Sense of Efficacy Scale* having almost an appropriate model or *goodness of fit* and being suitable for use in research models to measure the targeted construct, the low fit of the RMSEA value needs to be considered as a caution before using the scale in further studies. Upon further inspection of the model conducted using multigroup confirmatory factor analysis by gender, the relationship between *Class Management* and *Student Engagement* is more notable among female teachers, however, due to the lower numbers of fitness criteria (GFI: 0.995; CFI: 0.880; TLI: 0.845; RNI: 0.880; SRMR: 0.057; RMSEA: 0.111) of the model, further study is required. However, it should be noted that the separate unidimensional calibration of the Rasch model supports the construct validity of the Indonesian TSES. Thus, we conclude that the Indonesian TSES is ready for future studies.

Acknowledgment

The authors would like to thank all the teachers of children with special needs who were involved in this study. Our appreciation also extends to the inclusion schools and schools for special needs.

Conflict of Interest

The authors declare no competing interests.

Authors Contribution

RSR is responsible for research concept and design, data collection, data analysis and interpretation, and article writing. MC supervision and final approval of the article. KB supervision and final approval of the article. HP supervision and final approval of the article.

References

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), Article 1. https://doi.org/10.1037/amp0000191

Arias, C. R., Calago, C. N. S., Calungsod, H. F. B., Delica, M. A., Fullo, M. E., & Cabanilla, A. J. B. (2023). Challenges and Implementation of Inclusive Education in Selected Asian Countries: A Meta-Synthesis. *International Journal of Research in Education and Science*, *9*(2), Article 2. https://doi.org/10.46328/ijres.3089

Azwar, S. (2019). *Penyusunan Skala Psikologi*. Pustaka Pelajar.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), Article 2. https://doi.org/10.1037/0033-295X.84.2.191

Bond, T. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences, Third Edition* (3rd ed.). Routledge. https://doi.org/10.4324/9781315814698

Chachamovich, E., Fleck, M. P., Trentini, C. M., Laidlaw, K., & Power, M. J. (2008). Development and validation of the Brazilian version of the Attitudes to Aging Questionnaire (AAQ): An example of merging classical psychometric theory and the Rasch measurement model. *Health and Quality of Life Outcomes*, *6*(1), Article 1. https://doi.org/10.1186/1477-7525-6-5

Christensen, K. B., Makransky, G., & Horton, M. (2025, April 23). *Critical Values for Yen's Q3: Identification of Local Dependence in the Rasch Model Using Residual Correlations*. Applied Psychological Measurement. https://journals.sagepub.com/doi/10.1177/0146621616677520

Fabelico, F., & Afalla, B. (2020). Perseverance And Passion In The Teaching Profession: Teachers' Grit, Self-Efficacy, Burnout, And Performance. *Journal of Critical Reviews*, *7*(11), Article 11. https://doi.org/10.31838/jcr.07.11.17

Faragher, R., Robertson, P., & Bird, G. (2020). *International guidelines for the education of learners with Down syndrome | VOCEDplus, the international tertiary education and research database*. https://www.voced.edu.au/content/ngv%3A89128

Gálvez-Nieto, J. L., Salvo-Garrido, S., Domínguez-Lara, S., Polanco-Levicán, K., & Mieres-Chacaltana, M. (2023). Psychometric properties of the Teachers' Sense of Efficacy Scale in a sample of Chilean public school teachers. *Frontiers in Psychology*, *14*, 1272548. https://doi.org/10.3389/fpsyg.2023.1272548

Handayani, T., & Rahadian, A. S. (2013). Peraturan Perundangan Dan Implementasi Pendidikan Inklusif. *Masyarakat Indonesia*, *39*(1), Article 1. https://doi.org/10.14203/jmi.v39i1.307

Hayat, B., Putra, M. D. K., & Suryadi, B. (2020). Comparing item parameter estimates and fit statistics of the Rasch model from three different traditions. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *24*(1), Article 1. https://doi.org/10.21831/pep.v24i1.29871

Holster, T. A., & and Lake, J. (2016). Guessing and the Rasch Model. *Language Assessment Quarterly*, *13*(2), Article 2. https://doi.org/10.1080/15434303.2016.1160096

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), Article 4. https://doi.org/10.1037/1082-989X.3.4.424

Klassen, R. M., & Chiu, M. M. (2011). The occupational commitment and intention to quit of practicing and pre-service teachers: Influence of self-efficacy, job stress, and teaching context. *Contemporary Educational Psychology*, *36*(2), Article 2. https://doi.org/10.1016/j.cedpsych.2011.01.002

Kleinsasser, R. C. (2014). Teacher efficacy in *Teaching and Teacher Education*. *Teaching and Teacher Education*, *44*, 168–179. https://doi.org/10.1016/j.tate.2014.07.007

Luong, R., & Flake, J. K. (2023). Measurement invariance testing using confirmatory factor analysis and alignment optimisation: A tutorial for transparent analysis planning and reporting. *Psychological Methods*, *28*(4), Article 4. https://doi.org/10.1037/met0000441

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), Article 2. https://doi.org/10.1037/1082-989X.1.2.130

Maulana, H., Rangkuti, A. A., Sumintono, B., & Utami, L. D. (2020). Testing of the Indonesian Version of the Instrument "Teachers' Sense of Efficacy Scale" Using Rasch Modelling [Pengujian Kualitas Instrumen Teachers' Sense of Efficacy Scale Versi Bahasa Indonesia Menggunakan Pemodelan Rasch]. *ANIMA Indonesian Psychological Journal*, *35*(2), Article 2. https://doi.org/10.24123/aipj.v35i2.2905

Noviandari, H., & Huda, T. F. (2018). Peran Sekolah Dalam Pendidikan Anak Berkebutuhan Khusus Di Sdlb Pgri Bangorejo Banyuwangi. *Jurnal Psikologi : Jurnal Ilmiah Fakultas Psikologi Universitas Yudharta Pasuruan*, *5*(1), Article 1.

Pichardo, M. C., Cano, F., Garzón-Umerenkova, A., de la Fuente, J., Peralta-Sánchez, F. J., & Amate-Romera, J. (2018). Self-Regulation Questionnaire (SRQ) in Spanish Adolescents: Factor Structure and Rasch Analysis. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01370

Putri, F. A. R., & Fakhruddiana, F. (2018). Self-efficacy guru kelas dalam membimbing siswa slow learner. *JPK (Jurnal Pendidikan Khusus)*, *14*(1), Article 1. https://doi.org/10.21831/jpk.v14i1.25161

Rahayu, W., Hayat, B., & Putra, M. D. K. (2024). Analisis rasch: Aplikasi dan interpretasi. *UNJ Press*.

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), Article 3. https://doi.org/10.1037/0022-0663.99.3.611

Suseno, M., Hayat ,Bahrul, Putra ,Muhammad Dwirifqi Kharisma, Bien ,Jonna Karla, Rachmawati ,R., & and Hartanto, H. (2022). A differential item functioning (DIF) analysis of the mobile phone problem use scale in Indonesian schools with and without smartphone banned policy. *Cogent Psychology*, *9*(1), Article 1. https://doi.org/10.1080/23311908.2022.2137306

Tarnoto, N. (2016). Permasalahan-Permasalahan Yang Dihadapi Sekolah Penyelenggara Pendidikan Inklusi Pada Tingkat Sd. *Humanitas: Indonesian Psychological Journal*, *13*(1), Article 1. https://doi.org/10.26555/humanitas.v13i1.3843

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), Article 7. https://doi.org/10.1016/S0742-051X(01)00036-1

Umar, J., & Nisa, Y. F. (2020). Uji Validitas Konstruk dengan CFA dan Pelaporannya. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)*, *9*(2), Article 2.

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, *33*(3), Article 3. https://doi.org/10.6018/analesps.33.3.268401

Wright, B. D. (1996). Local dependency, correlations and principal components. *Rasch Measurement Transactions*, *10(3)*, 509–511.

Zee, M., & Koomen, H. M. Y. (2016). Teacher Self-Efficacy and Its Effects on Classroom Processes, Student Academic Adjustment, and Teacher Well-Being: A Synthesis of 40 Years of Research. *Review of Educational Research*, *86*(4), Article 4.