Journal of Psychology and Instruction
Volume 9, Number 1, 2025 pp. 27-42
P-ISSN: 2597-8616, E-ISSN : 2549-4589
Open Access: https://doi.org/10.23887/jpai.v9i1.49814

Psychometric Properties and Latent Subclass Analysis Mixture Rasch Model of on Self-Adjustment Questionnaire

Merinta Wira Ababiel[1*], Ananta Yudiarso[2]

[1,2] Psychology, Surabaya University, Surabaya, Indonesia

ARTICLE INFO

Article history:
Received January 10, 2025
Accepted March 30, 2025
Available online April 25, 2025

Kata Kunci:
Analisis Latent Class, Mixture Rasch Model, *Penyesuaian Diri*

Keywords:
Latent Class Analysis, Mixture Rasch Model, Self-Adjustment

ABSTRAK

Penelitian mengenai validasi pengukuran penyesuaian diri menggunakan Model Rasch masih terbatas di Indonesia, terutama penelitian yang berfokus pada mahasiswa. Penelitian ini bertujuan untuk menilai sifat psikometrik angket penyesuaian diri dengan menggunakan pendekatan kuantitatif dan analisis Mixture Rasch Model. Kuesioner penyesuaian diri dikembangkan berdasarkan teori Haber & Ruyon (1984). Sebanyak 217 mahasiswa berpartisipasi dalam penelitian ini. Hasilnya menunjukkan bahwa asumsi unidimensi terpenuhi setelah penghapusan satu item dari setiap subskala: ACS_6, PSI_12R, GIR_19R. Skala Likert lima poin berfungsi secara efektif, memungkinkan peserta untuk membedakan berbagai pilihan respons. Reliabilitas item untuk sebagian besar subskala tercapai, berkisar antara 0,70 hingga 0,90, kecuali untuk subskala PSI (0,36). Namun pemisahan orang untuk semua sub-skala tidak memenuhi kriteria, berkisar antara 0,70 dan 1,86 (≤ 3). Seluruh item dalam penelitian ini tidak menunjukkan DIF dan tidak ada subkelas laten berdasarkan hasil analisis Mixture Rasch. Secara keseluruhan, kuesioner penyesuaian diri menunjukkan kinerja yang cukup baik, kecuali untuk keandalan dan keterpisahan orang, yang tidak memenuhi kriteria yang dipersyaratkan. Sebagai keterbatasan penelitian, item tersebut kurang efektif dalam menjaring peserta dengan kemampuan rendah dan tinggi.

ABSTRACT

Research on the validation of self-adjustment measurement using the Rasch Model remains limited in Indonesia, especially studies that focus on university student. This study aimed to assess the psychometric properties of the self-adjustment questionnaire using a quantitative approach and Mixture Rasch Model analysis. The self-adjustment questionnaire was developed based on the theory of Haber & Ruyon (1984). A total of 217 university students participated in the study. The results indicated that the unidimensionality assumption was met after the removal of one item from each sub-scale: ACS_6, PSI_12R, GIR_19R. The five point Likert scale functioned effectively, allowing participants to differentiate among various response options. Item reliability for most sub-scales was achieved, ranging from 0.70 to 0.90, except for the PSI sub-scale (0.36). However, person separation for all sub-scales did not meet the criteria, ranging between 0.70 and 1.86 (≤ 3). All items in this study exhibited no DIF and no latent subclass based on the Mixture Rasch analysis results. Overall, the self-adjustment questionnaire demonstrated fairly good performance, except for person reliability and separation, which did not meet the required criteria. As a limitation of the study, the items were less effective in capturing participants with low and high abilities.

1. INTRODUCTION

Self-adjustment is an essential aspect of university life for students. Self-adjustment defined as a process of how well an individual is able to cope with constantly changing situations and conditions, in which they feel aligned with their environment and experience satisfaction in fulfilling their needs (Haber as cited in (Ababiel et al., 2023). As widely acknowledged, higher education at the university level introduces a unique dynamic that differs significantly from earlier educational stages, such as Kindergarten (TK), Elementary School (SD), Junior High School (SMP), and Senior High School (SMA) or Vocational High School (SMK). These distinctions encompass various aspects, including differences in curriculum structure, learning systems, academic workloads, individual and academic responsibilities, levels of discipline, and the nature of social interactions that students experience throughout their university journey (Daulay, 2021).

Corresponding author
*E-mail addresses: merintawiraa@gmail.com (Merinta Ababiel)

According to Fadli as cited in (Wistarini, 2019), the curriculum at the higher education level (university) differs significantly from that of primary and secondary education. While primary and secondary education focus more on fostering students' humanistic aspects, such as character development and general knowledge, the higher education curriculum is designed to emphasize scientific advancement, mastery of specific disciplines, and preparation for the workforce. Unlike the earlier educational stages, where the primary goal is holistic personal growth, higher education places greater emphasis on specialized expertise, professional skills, and career readiness to equip students for their future professions (Direktorat Jenderal Pendidikan Tinggi, 2024).

Among all educational transitions, the shift from high school to university represents the most critical and demanding phase (Wistarini & Marheni, 2019). In this context, critical and demanding phase refers to the complexity of academic demands, which makes university students vulnerable to emotional challenges. These challenges may include feelings of anxiety, fear, and confusion about what actions to take and whom to ask for guidance when facing uncertainties. Additionally, students may struggle to build social relationships, experience low academic performance, lose motivation, or even fail to complete their studies (Agestia et al., 2024). Furthermore, research conducted by Saalino et al., (2022) on the self-adjustment of first-year university students indicates that more than 50% of students encounter challenges in adapting to university life. These challenges include difficulties in accessing study materials outside the campus, struggles in establishing social connections within the university environment, a lack of confidence in seeking clarification from lecturers regarding unclear course materials, and issues related to time management in completing assigned tasks. Another study by Agestia et al., (2024) found that 95% of students experience culture shock when entering university life, leading to feelings of anxiety, fear, lack of confidence, loneliness, and health issues. Furthermore, students struggle to adapt academically, socially, emotionally, and institutionally. Various issues in self-adjustment can have negative effects on students, leading to academic stress. This, in turn, may result in a decline in academic motivation, which subsequently leads to decreased academic performance and an increased risk of student dropout (Pascoe et al., 2020).

Based on the above discussion, validating the Self-Adjustment Questionnaire Measurement is a critical step in ensuring its effectiveness and accuracy. Thus far, validation using the Mixture Rasch Model has not been conducted for the Self-Adjustment Questionnaire in Indonesia, and there is still limited research on self-adjustment with a specific focus on university students. Therefore, this study employs the Mixture Rasch Model to assess the validity of the Self-Adjustment Questionnaire. This questionnaire was developed by the researcher based on the five key aspects of self-adjustment by Haber & Runyon (1984): (1) Accurate in perceiving reality. Individuals carefully set realistic goals for their future, align their objectives with their abilities, and evaluate various consequences of their actions; (2) Ability to manage stress and anxiety. Individuals can leverage painful experiences to overcome feelings of stress and anxiety; (3) A Positive self-image. This involves an individual's assessment of their positive self-concept, which can be evaluated both personally and through others' perceptions; (4) Ability to express emotions. Individuals demonstrate good emotional expression and control; (5) Good interpersonal relationships. Well-adjusted individuals can establish meaningful and beneficial social connections. Recent studies highlight that self-adjustment has a strong correlation with individual success across various contexts, such as education, work, and social interactions (Arifin, 2013; Mamahit & Defiani, 2022; Sasti et al., 2023).

In the context of psychometrics, measurement tools used in research must possess high validity, reliability, and sensitivity to individual differences (Yulisharyasti et al., 2023). The Standards for Educational and Psychological Testing define validity as the extent to which empirical evidence, supported by strong theoretical reasoning and logical analysis, sufficiently justifies and thoroughly substantiates the interpretation of test scores. This validation ensures that the test scores align with their intended purpose, application, and broader contextual use across various settings. Furthermore, internal structure validity can be assessed using sophisticated tools and methodologies, including the integration of Information and Communication Technology (AERA et al., 2014). One effective approach for assessing the internal structure of psychological instruments is through Rasch Model, whichprovides a robust analytical framework, thoroughly examining the interactions between raters and items, contributing to a more comprehensive assessment of the validity and reliability of measurement instruments. Additionally, it ensures that test items function cohesively, accurately capturing the intended construct in a structured and meaningful way (Avinç & Doğan, 2024; Rios & Wells, 2014).

The Rasch Model ensures the quality of a measurement tool by confirming its internal validity through an item-response theory approach to analyze data from tests or questionnaires (Bond & Fox, 2021). Because, the test instrument that has not undergone item quality analysis leads to a quasi-assessment, affecting the accurate measurement of a student's actual ability (Fauziana & Wulansari, 2021). The Rasch Model measures latent traits (such as abilities, attitudes, or psychological characteristics) based on individuals' responses to items. The Rasch model offers several advantages: (1) interval measurement. It

produces equal interval units, known as logits, when the data align with the model, facilitating criterion-referenced interpretations through qualitative descriptions of the examinee's agrees or does not agree to do; (2) separation of parameters. It separates person ability from item difficulty, meaning estimates are not influenced by the specific sample or test; (3) validity and reliability assessment. It offers comprehensive diagnostic tools, including fit statistics, error estimates, and variable maps, to thoroughly evaluate the validity and reliability of item and person measures; (4) handling missing data. It allows for the inclusion of missing data without compromising the overall measurement quality; (5) criterion-referenced and norm-referenced data are presented, providing person measures and item calibrations on a unified linear scale that enhances the ability to make meaningful comparisons; (6) adjustments for systematic variations. The model can account for variations due to judge severity, item difficulty, or testing conditions (Smith et al., 2002; Araya et al., 2018).

The challenges frequently encountered in Rasch Model analysis involve heterogeneity within the population, which can influence measurement outcomes. Previous studies often overlooked this aspect of heterogeneity, particularly in the context of self-adjustment measurement. Jannah, (2013) examined the validity of a self-adjustment measurement tool and found that all items were unidimensional, measuring a single factor. However, this study was limited to a specific population students in an Islamic boarding school. This limitation creates a gap, as the validity results cannot be generalized to other populations with differing characteristics, such as private school students or public school students. Credé & Niehorster, (2012) conducted a detailed and comprehensive study with the objective of thoroughly evaluating the validity andreliability of theSACQ (Student Adaptation College Questionnaire) in relation tofirst-year college students' academic performance, psychological well-being, social adjustment, and overall adaptation to the challenges of the college environment. The findings conclusively demonstrated that the SACQ exhibited an exceptionally strong degree of internal consistency (Cronbach's alpha = 0.89), which highlights its reliable, robust, and consistent measurement properties across diverse contexts. However, the generalizability of the findings to non-student populations or advanced-level students has not been comprehensively tested. As stated by Che Lah et al., (2022), Rasch Model analysis enables the classification of questionnaire items into different levels of difficulty, such as easy, moderate, and difficult. Moreover, it has the capability to transform ordinal data into ratio data, which allows for more appropriate statistical analyses, as ratio data represents the highest level of measurement in statistical evaluation. Additionally, Rasch Model analysis enhances the overall effectiveness of questionnaires in research by facilitating the identification of duplicate items and scales within a questionnaire, validating Cronbach's alpha scores, and assessing both item and person reliability. As a result, Rasch Model analysis is widely recognized as a powerful tool for ensuring the reliability and validity of questionnaires.

Moreover, a systematic review by Zhang & Goodson, (2011) analyzed data from 64 studies that explored measurement tools in various research on the self-adjustment of international students. Studies employing tools such as the Brief Symptom Inventory (BSI) for psychological adjustment and the Sociocultural Adaptation Scale (SCAS) for sociocultural adjustment revealed specific limitations. The BSI primarily focuses on negative psychological symptoms (distress) without addressing positive aspects, such as effective coping strategies or psychological resilience. In contrast, the SCAS emphasizes measuring social and cultural adjustment but does not encompass emotional and psychological dimensions. These limitations render both tools less comprehensive for holistically measuring psychological adjustment. In fact, psychological and emotional dimensions must be considered to understand how these factors influence respondents' ability to answer questionnaire items. According to Nur et al., (2022), the Rasch Model serves as an analytical tool designed to assess both respondents' abilities and item difficulty by adjusting for the challenges faced by respondents. Rasch Model analysis provides insights into the construct validity of an instrument, ensuring that the collected data is accurate and reliable.

Another study by Nyimas & Rulanggi, (2022)conducted a detailed and comprehensive study with the objective of thoroughly evaluating the validityand reliabilityof the SACQ (Student Adaptation College Questionnaire) in relation to first-year college students using Confirmatory Analysis Factor (CFA) method. The validity test results of the Student Adaptation to College Questionnaire (SACQ) indicate that the item discrimination indices rangefrom 0.149 to 0.593, witha Cronbach's alpha coefficient of 0.873. However, the Confirmatory Factor Analysis (CFA) used in this study is limited in detecting whether all 40 items exhibit bias toward specific groups within the student population. In contrast, the Rasch Model allows for a more precise Differential Item Functioning (DIF) analysis, enabling the identification of whether certain items are systematically easier or more difficult for groups with specific characteristics, such as gender or cultural background (Wahyuni, 2022).

Based on the validity gap of the self-adjustment measurement tool from the study above, Wang et al., (2018) highlighted that employing the Mixture Rasch Model is effective in uncovering latent subgroups within a population, thus providing deeper insights into different response patterns. This study also verify

self adjustment scale with the Mixture Rasch Model analyze. Gluck & Spiel, (1997) provided a detailed description of the Mixture Rasch Model, characterizing it as an advanced statistical approach that seamlessly integrates the principles of latent class analysis with the robust framework of the Rasch Model. This innovative method allows researchers to simultaneously identify unobserved subgroups within a population and assess item response patterns, thereby offering deeper insights into individual differences and latent traits across diverse contexts. The fundamentaliidea is that the Rasch model fails to align effectively with the theoretical expectations of the analyzed data, this misalignment may stem from the inherent heterogeneity present within the sample. In other words, there might be two or more subgroups (classes) within the sample, which requires using the perspective of a mixture distribution model. MRM can identify the most significant differences between item parameters across the compared sample. Thus, it eliminates the need to manually identify and separate the sample based on its characteristics to find constant item parameters for all individuals. Another advantage of using MRM is its ability to test group differences in applying different strategies to respond to test items. Since the use of different strategies is related to item difficulty levels, this type of test analysis requires a model that accounts for varying item parameters across subpopulations Rost & Davier (as cited in Widhiarso, 2010).

The Mixture Rasch Model (MRM) differs from Latent Class Analysis (LCA) by integrating the Rasch Model with LCA to uncover latent subgroups within the data that display distinct response patterns. In MRM, it is assumed that respondents belong to multiple latent populations, with the Rasch Model applied separately to each latent class, each having its own unique item parameters (Jiao et al., 2011). The main difference between MRM and LCA lies in their approach to data. LCA focuses on identifying latent subgroups based on categorical response patterns, without considering individual ability levels. In contrast, MRM accounts for both qualitative and quantitative differences by integrating latent class analysis and item response modeling. This allows the identification of subgroups with differences in strategies or abilities (Toker & Green, 2021). MRM is particularly useful in situations where the assumption of population homogeneity is violated, as it allows researchers to identify heterogeneity among participants, uncover associated item structures, determines latent class sizes, and analyze latent score distributions. Thus, MRM provides deeper insights into the data structure compared to traditional LCA or Rasch Models (Toker & Seıdel, 2023). LCA categorizes a population into distinct latent groups based on similar response patterns, whereas MRM addresses heterogeneity in Rasch data by identifying subpopulations with varying Rasch parameters, such as item difficulties (Mccreary et al., 2013).

The application of the Mixture Rasch Model (MRM) in psychometric data analysis is deeply rooted in evidence-based principles, as it rigorously examines the internal structure and dimensionality of an instrument. This approach aligns with the high standards set forth by the International Test Commission and adheres to the comprehensive guidelines established by the American Educational Research Association (AERA et al., 2014), ensuring that the analyses are both methodologically robust and theoretically sound for diverse research and assessment contexts. These guidelines highlight the necessity of ensuring that measurement instruments possess valid and reliable internal structures while considering heterogeneity within the user population. Unlike Differential Item Functioning (DIF), which requires prior knowledge of groups, MRM identifies latent subpopulations in the data, focusing on population heterogeneity (Rost, 1990). Conversely, DIF is utilized to assess item bias towards a specific known group, such as gender or culture, by comparing item responses after controlling for ability (Khine, 2021). MRM allows for deeper analysis by identifying latent subgroups within the population that may have different response parameters (Vermunt, 2022; Vermunt & Magidson, 2002). This is relevant to the ITC and AERA recommendations to test whether items on an instrument function consistently across groups. MRM provides empirical evidence of the existence of subpopulations with different latent structures that are not detected by standard Rasch models (Mccreary et al., 2013). MRM ensures that the measurement is not only appropriate for the population as a whole but also fair and accurate for each subgroup, in accordance with the principles of fairness and equity in psychological testing (AERA et al., 2014).

2. METHOD

This research design employs a psychometric validation framework combined with a quantitative approach, with the primary objective of thoroughly evaluating the self-adjustment questionnaire using Mixture Rasch Model to gain an in-depth and thorough insight into the questionnaire's structure and properties, with a focus on assessing exercise motivation. Additionally, it places a strong emphasis on assessing the instrument's internal validity, reliability, and overall effectiveness in capturing the intended constructs across diverse populations and contexts.

The sampling process in this study was carried out through the distribution of physical questionnaires directly to potential participants, ensuring accessibility and convenience for data collection.

The sampling technique utilized was non-random sampling, which explicitly signifies that not all individuals within the broader population were afforded an equal opportunity. This approach was deliberately implemented to secure participants that is highly relevant and aligned with the specific research objectives. It focused on selecting individuals who met clearly defined inclusion criteria, such as demographic characteristics, situational factors, or psychological attributes, deemed essential for addressing the study's central questions and achieving its intended outcomes. The type of non-random sampling used was the saturated sampling technique, where all members of the population meeting specific criteria are included in the study. The criteria referred to in this study include all students from the 2021 and 2022 faculty class, totaling 217 students. All participants have provided informed consent and agreed to take part in the study, which was conducted offline in their respective classes. By including the entire population, saturated sampling minimizes selection bias and provides a comprehensive, representative dataset of the target group (Creswell, 2018).

The principle of instrument calibration significantly affects the sample size. When an instrument is calibrated on a sample that does not overlap with the participants, slight variations in results are expected. Consequently, a small sample size can lead to unstable calibration outcomes and reduced sensitivity in capturing accurate results. On the other hand, using a larger sample size can minimize differences in calibration results, though it requires careful consideration of cost and time efficiency. According to Linacre (1994), for Rasch analysis with a 99% confidence level, the sample size that can be used ranges from 108 to 243 respondents. This study using the Raosoft sample size calculator feature with a marginof error of 5%and a confidence levelof 86.1%, which involved 217 participants, including 13 females and 204 males, all of whom are students of one of the Universities in Surabaya.

This research utilizes Rasch analysis and the Mixture Rasch Model as its core methodologies, employing the advanced analytical tools WINSTEP® 5.1.0 and Jamovi 2.3.28 software for data processing and evaluation. The psychometric properties assessed encompass a range of critical dimensions, including the overall reliability of the instrument, the reliability of individual respondents, the consistency of items, and the unidimensionalityof the measurement scale. Additionally, the study evaluates the functionality and appropriateness of the rating scale and conducts an in-depth analysis of potential bias using Differential Item Functioning (DIF) techniques, ensuring that the instrument performs equitably across diverse subgroups within the sample.

The data collection process in this study employed a self-adjustment scale specifically designed based on the theoretical framework by Haber & Runyon (1984). This scale consists of 19 constructed items that reflect the key aspects of self-adjustment. Each item is rated using a 5-point Likertscale with response optionsfrom: 1 ("strongly disagree"), 2 ("disagree"), 3 ("neutral"), 4 ("agree"), to 5 ("strongly agree"). The detailed blueprint of the scale, outlining its structure and alignment with the theoretical aspects, is provided below for reference:

Tabel 1. The Self-Adjustment Questionnaire Blueprint

Sub-Scale	Item Coding	Item
Accurate perception of reality	APR_1	Saya bertanggungjawab atas semua keputusan yangsaya ambil.
	APR_2	Saya yakin bisa mencapai cita-cita dengan kemampuan saya saatini.
	APR_3R	Saya merasa jurusan saat ini kurang sesuai dengankeinginan
	APR_4R	Saya merasa pesimis bisa lulus dengan nilai yangbagus.
Ability to Cope with Stress and Anxiety	ACS_5	Ketika berbeda pendapat dengan teman, saya memilih untuk menghindari berselisihpaham.
	ACS_6	Saya berpikir bahwa kegagalan merupakan pelajaran yang berharga agar saya menjadi lebihbaik lagi.
	ACS_7R	Saya merasa tertekan dengan model pembelajaran yang diterapkansaat ini.
	ACS_8R	Saya kesulitan berpikir jernih ketika menghadapimasalah yang berat.
A Positive Self-Image	PSI_9	Saya merasa bisa sukses di masa depan dengan kemampuan sayasaat ini.
	PSI_10	Saya bertanya pada teman ketika kesulitan mengerjakan tugaskuliah.
	PSI_11R	Saya merasa kemampuan teman-teman jauhdiatas kemampuan saya.
	PSI_12R	Saya pesimis dengan kemampuan yang sayamiliki.
Ability to Express Feelings	AEF_13	Lebih baik bertukar pikiran dengan orang lain ketika menghadapi masalah yangsulit.
	AEF_14	Saya merasa kesulitan berinteraksi denganorang yang baru dikenal.
	AEF_15R	Saya merasa malu jika menceritakan masalah padateman.

Sub-Scale	Item Coding	Item
GoodInterpersonal Relations	AEF_16R	Saya lebih memilih untuk mengerjakantugas kelompokk sendiri daripada bersama temann.
	GIR_17	Saat ada konflik dengan teman, saya dapat menyelesaikannya denganbaik.
	GIR_18	Sayasenang jika dapat mengerjakan tugas bersamateman-teman.
	GIR_19R	Saya tidak mampu bekerjasama dengan siapapun saat mendapatkan tugaskelompok.
	GIR_20R	Saya menolak bila diajakkumpul-kumpul dengan teman.

3. RESULT AND DISCUSSION

3.1 Unidimensionality

The unidimensionality of each sub-scale was assessed based on the raw variance explained by the measure, which served as the primary criterion in the Rasch Principal Component Analysis of Residuals (PCAR). Holster & Lake, (2016) emphasized that a raw variance value exceeding 40% provides sufficient evidence supporting the presence of unidimensionality within the scale. Additionally, the eigenvalue of the first contrast should remain below 2.0, as a lower value suggests that the residuals mostly consist of random noise rather than indicating the presence of an additional dimension within the data. This requirement is essential to ensure that the measurement remains unidimensional, allowing it to effectively assess the intended construct without interference from other unintended variables.

Tabel 2. Unidimensionality

Self-Adjustment Sub-Scale	Raw Variance Explained by Measure	First Contrast	After Item Removal	After Item Removal	Removal Item	Conclusion
Accurate Perception of Reality	57.8%	1.5	-	-	-	Undimensional
Ability to Cope with Stressand Anxiety	36.5%	0.39	40.7%	1.7	ACS_6	Undimensional
A Positive Self-Image	36.6%	1.6	44.6%	1.7	PSI_12R	Undimensional
Ability to Express Feelings	62.6%	1.5	-	-	-	Undimensional
Good Interpersonal Relations	36.4%	0.41	43.4%	1.6	GIR_19R	Undimensional

The dimensionality analysis results reveal that in the initial evaluation of the subscales Ability to Cope with Stress and Anxiety, A Positive Self-Image, and Good Interpersonal Relations, the raw variance explained by the measures was found to be below 40%. This finding suggests that these subscales measure more than one underlying latent trait, indicating their multidimensional nature rather than capturing a single unified construct.

A second analysis was conducted by removing one item from each of the two subscales: ACS_6 “Saya berpikir bahwa kegagalan merupakan pelajaran yang berharga agar saya menjadi lebihbaik lagi”, PSI_12R “Saya pesimis dengan kemampuan yang sayaa miliki” and GIR_19R “Saya tidak mampu bekerjasama dengan siapapun saat mendapatkan tugaskelompok”, resulting in raw variance explaineddby measures exceeding 40% for all three subscales. This indicates that the items consistently measure a single latent trait, thereby meeting the criteria for unidimensionality. Theinterpretation of item unidimensionality, as measured by the raw variancee explained, is interpreted as follows: a value exceeding 20% is deemed acceptable, over 40% is regarded as good, and surpassing 60% is classified as excellent (Muntazhimah et al., 2020).

The unidimensionality results shows that in the first analysis of the subscales Ability to Cope with Stressand Anxiety item ACS_6, A Positive Self-Image item PSI_12R, and Good Interpersonal Relations item GIR_19R assess more than one latent trait and can thus be considered multidimensional. Then, a second

analysis was conducted by removing one item from each of the three subscales: ACS_6 "Saya berpikir bahwa kegagalan merupakan pelajaran yang berharga agar saya menjadi lebihbaik lagi", PSI_12R "Saya pesimis dengan kemampuan yang sayamiliki" and GIR_19R "Saya tidak mampu bekerjasama dengan siapapun saat mendapatkan tugass kelompok", resulting in raw variance explaineddby measures exceeding 40% for all three subscales, indicates that the items consistently measure a single latent trait, thereby meeting the criteria for unidimensionality.

Based on previous research conducted by Nyimas & Rulanggi, (2022) on the analysis of the Student Adaptation to College Questionnaire (SACQ) as a measurement instrument for college adjustment among first-year students using Confirmatory Factor Analysis (CFA), the findings indicate that CFA was primarily used to confirm that the four aspects of SACQ, comprising 40 items, measure the theoretical construct adapted from Soraya (2020). However, the study did not further explore whether all items genuinely reflect their respective aspects without the influence of additional factors. Meanwhile, based on the results of the previously mentioned dimensionality, it can be concluded that all five aspects of self-adjustment exhibit unidimensionality. This indicates that each test item is designed to measure only a single specific ability, ensuring the consistency and precision of the measurement within each aspect (Handayani et al., 2023).

3.2 Rating Scale Diagnostics

The measurement of self-adjustment in this study utilized a 5point Likert scale, which provided respondents with the following response: "strongly disagree", "disagree", "neutral", "agree", and "strongly agree". A diagnostic assessment using a rating scale was carried out to evaluate whether respondents were able to clearly distinguish between each available response option (Gunawan, 2023).

Tabel 3. Rating scale diagnostics

Self-Adjustment Sub-Scale	Likert	Observed Count	Observed Percentage	Andrich Threshold	Category Measures
AccuratePerception of Reality	1 Strongly Disagree	140	16	NONE	-3.11
	2 Disagree	226	26	-1.86	-1.27
	3 Neutral	221	25	-0.41	0.08
	4 Agree	159	18	0.67	1.30
	5 Strongly agree	122	14	1.60	2.93
Ability toCope with Stress and Anxiety	1 Strongly Disagree	107	16	NONE	-2.63
	2 Disagree	176	27	-1.32	-1.01
	3 Neutral	173	27	-0.27	0.04
	4 Agree	125	19	0.46	1.04
	5 Strongly agree	70	11	1.14	2.52
A PositiveSelf-Image	1 Strongly Disagree	105	16	NONE	-2.68
	2 Disagree	158	24	-1.34	-1.09
	3 Neutral	179	27	-0.45	0.00
	4 Agree	136	21	0.46	1.09
	5 Strongly agree	73	11	1.33	2.67
Ability toExpress Feelings	1 Strongly Disagree	158	18	NONE	-3.36
	2 Disagree	237	27	-2.16	-1.34
	3 Neutral	178	21	-0.23	0.09
	4 Agree	172	20	0.50	1.37
	5 Strongly agree	123	14	1.90	3.16
Good Interpersonal Relations	1 Strongly Disagree	104	16	NONE	-2.66
	2 Disagree	166	25	-1.36	-1.01
	3 Neutral	163	25	-0.22	0.05
	4 Agree	133	20	0.44	1.04
	5 Strongly agree	85	13	1.15	2.52

Based on theresults of rating scale diagnostics, the Andrich Threshold column showed values progressing from small (negative) to large (positive). This sequential order of thresholds confirms that the response categories were functioning as intended, allowing participants to discern meaningful distinctions between each level of the scale. The findings further validate the scale's effectiveness, indicating that it provides a reliable and precise tool for measuring self-adjustment by ensuring that respondents perceive the response options distinctly and respond accordingly to the items in alignment with their actual experiences and perspectives.

The self-adjustment measurement tool used in this test has a good level of item reliability, but it is less sensitive in capturing the diversity of participants' abilities because the person separation across all subscales showed values below 3 and for item separation, the subscales APR, ACS and PSI displayed values below 3. However, item reliability was established meanwhile the separation index did not meet the criteria. This could be attributed to the limited numberoof items. According to Setiyawan, (2020), the reliability (accuracy) of measurements generally improves as the number of items included in the tool increases. A larger item pool enables a more thorough construct assessment, improving result precision and consistency of the results. Conversely, a measurement tool with a limited number of items may yield outcomes that are inconsistent or insufficiently representative of the construct, potentially reducing its ability to detect and capture meaningful differences in abilities or traits among participants.

The rating scale indicate that the self-adjustment measurement tool designed in this study is capable of measuring what it is intended to measure, and participants are able to differentiate the scale and the response options within this scale: “strongly disagree”, “disagree”, “neutral”, “agree”, “strongly agree”.

A different result was found in the study conducted by Syifa et al., (2023), which examined a college adjustment measurement instrument developed based on the theory and aspects proposed by Baker & Siryk (1984) and later adapted by Sumomba (2023). The study concluded that college adjustment among first-year students living away from home falls into the high-level category. However, the study did not provide an in-depth analysis of whether the college adjustment instrument was truly comprehensible and capable of accurately measuring the actual self-adjustment abilities of students from different regions.

3.3 Person-Item Reliability

The Rasch Model evaluates the reliability of both individuals and items. In this context, the instrument's capability to distinguish respondents based on the specific variable being measured is known as person reliability, which reflects the consistency and precision of the measurement in categorizing individuals according to their traits or abilities (Dwiliesanti & Yudiarso, 2022). Based on Fisher (2007), the reliability of person and item measurement is deemed adequate when it falls within the range of 0.67 to 0.80, while a value below 0.67 is considered low. Additionally, the person separation index also serves as an indicator of reliability, as it is calculated using the logit scale instead of raw scores, ensuring a more precise measurement. Furthermore, Linacre (2024) states that a person separation value must be greater than 3.0 to be categorized as high, signifying that the instrument possesses a strong capability to differentiate respondents based on their varying levels of the measured trait.

Tabel 4. Person-item Realiability

Self Adjustment Sub-Scale	Code	Person Reliability	Person Separation	Item Reliability	Item Separation
Accurate Perception of Reality	APR	0.74	1.68	0.75	1.71
Ability to Cope withStress and Anxiety	ACS	0.33	0.70	0.82	2.17
A Positive Self-Image	PSI	0.46	0.92	0.36	0.76
Ability to Express Feelings	AEF	0.78	1.86	0.91	3.18
Good Interpersonal Relations	GIR	0.44	0.88	0.91	3.19

In this study, the results show that participants were still inconsistent, as only the subscales APR and AEF demonstrated person reliability above 0.6. Person separation across all subscales also showed values below 3, indicating that theitems were not sensitive enoughto capture the diversity of participants' abilities. For item separation, the subscales APR, ACS and PSI displayed values below 3, suggesting that most participants lacked sufficient variability to identify item diversity. Additionally, almost all aspects demonstrated excellent item reliability (>0.7), except for the PSI sub-scale, which did not meet the criteria for good item reliability, with a score of 0.36.

Based on previous research by Safitri, (2020), which employed Classical Test Theory (CTT) to assess the reliability of the measurement instrument, a Cronbach's Alpha value of 0.824 was obtained, indicating high reliability. However, the study lacked a comprehensive explanation regarding whether the instrument remains suitable for measuring conditions when applied to different groups of students. This omission ultimately renders the comparison between items and respondents unreliable and invalid.. In line with Setiawati et al., (2023), the Rasch Model seems to provide a viable solution to the limitations present in various analytical methods currently utilized, including Classical Test Theory (CTT), which is widely applied in social science research, particularly in psychology. One significant drawback of CTT is that when different tests are employed to assess the same construct across various student populations, the validity of comparisons between items and respondents is compromised, making the results less reliable or invalid.

3.4 Item Fit

Item Fit serves as an indicator to evaluate whether items align with the model and function properly in measurement. The criteria for identifying misfitting items in this study were established based on rigorous psychometric guidelines. These criteria include: (1) an acceptable outfit mean square (MNSQ) value, which should fall within the range of 0.5 to 1.5, as recommended by Linacre (2002), ensuring that the item responses align well with the expected model; and (2) an acceptable outfit standardized (ZSTD) value, which should range between -2.0 and 2.0, as outlined by Sumintono & Widhiarso, (2015), to confirm that the items exhibit minimal deviation from the model's expectations (Sumintono, 2016).

Tabel 5. Item Fit Rasch Model

Sub-Scale	Item Coding	Question	Measure	Infit		Outfit		Pt. Measure
				MNSQ	ZSTD	MNSQ	ZSTD	
Accurate Perception ofReality	APR_1	Saya bertanggungjawab atas semua keputusan yang saya ambil.	-0.19	1.09	0.9	1.06	0.7	0.78
	APR_2	Saya yakin bisa mencapai cita-cita dengan kemampuan saya saat ini.	-0.03	0.87	-1.3	0.86	-1.5	0.79
	APR_3R	Saya merasa jurusan saat ini kurang sesuai dengankeinginan	-0.06	0.97	-0.3	0.98	-0.1	0.74
	APR_4R	Saya merasa pesimis bisa lulus dengan nilai yangbagus.	0.27	1.03	0.3	1.04	0.4	0.75
Ability to Copewith Stress and Anxiety	ACS_5	Ketika berbeda pendapat dengan teman, saya memilih untuk menghindari berselisihpaham.	-0.18	1.16	1.7	1.14	1.5	0.61
	ACS_7R	Saya berpikir bahwa kegagalan merupakan pelajaran yang berharga agar saya menjadi lebihbaik lagi.	-0.06	1.05	0.5	1.02	0.2	0.65
	ACS_8R	Saya merasa tertekan dengan model pembelajaran yang diterapkansaat ini.	0.24	0.79	-2.4	0.83	-1.9	0.62
A PositiveSelf-Image	PSI_9	Saya kesulitan berpikir jernih ketika menghadapimasalah yang berat.	0.04	0.80	-2.3	0.79	-2.4	0.69
	PSI_10R	Saya merasa bisa sukses di masa depan dengan	0.10	1.00	0.1	0.99	0.0	0.71

Sub-Scale	Item Coding	Question	Measure	Infit		Outfit		Pt. Measure
				MNSQ	ZSTD	MNSQ	ZSTD	
		kemampuan saya saat ini.						
	PSI_11R	Saya bertanya pada teman ketika kesulitan mengerjakan tugas kuliah.	-0.13	1.19	2.0	1.16	1.7	0.60
Ability to Express Feelings	AEF_13	Saya merasa kemampuan teman-teman jauh diatas kemampuan saya.	-0.27	1.00	0.1	1.08	0.8	0.78
	AEF_14	Saya pesimis dengan kemampuan yang sayamiliki.	0.47	1.10	1.0	1.06	0.6	0.78
	AEF_15R	Lebih baik bertukar pikiran dengan orang lain ketika menghadapi masalah yang sulit.	-0.26	1.03	0.4	1.04	0.4	0.81
	AEF_16R	Saya merasa kesulitan berinteraksi dengan orang yang baru dikenal.	0.06	0.82	-1.9	0.78	-2.4	0.83
Good Interpersonal Relations	GIR_17	Saya merasa malu jika menceritakan masalah padateman.	-0.28	0.95	-0.5	0.93	-0.8	0.70
	GIR_18	Saya lebih memilih untuk mengerjakan tugas kelompokk sendiri daripada bersama temann.	-0.04	0.90	-1.2	0.88	-1.3	0.67
	GIR_20R	Saat ada konflik dengan teman, saya dapat menyelesaikannya dengan baik.	0.33	1.17	1.8	1.15	1.5	0.58

The analysis results reveal that all items in the self-adjustment measurement meet the established criteria for fit, with mean square (MNSQ) values consistently falling within the acceptable range of 0.5 to 1.5, while the standardized (ZSTD) values range between -2 and 2. This indicates that each item aligns well with the measurement model and exhibits strong consistency. Furthermore, the PMC (Point Measure Correlation) values for all itemsare above 0.4, which is categorized as good based on the standards outlined by Tanujaya & Yudiarso, (2023). Overall, all items in the self-adjustment measurement tool are appropriate for use without the need to remove any items.

Based on the item accuracy criteria with the model that examines the normal function in measurement, in general, when viewed from the values of infit, outfit mean square, and standardized, as well as point measure correlation, all meet the requirements. Therefore, the measurement tool is considered to be a fit according to the Rasch Model and suitable for use without the need to discard any items. In line with the findings in study by Gunawan, (2023), with analysis results of infit,outfit,point measure correlation that meet therequirements indicate that the measurementtool is suitablefor use.

Based on previous research by Bahiyah & Savitri, (2018), which also conducted a validation test on a self-adjustment measurement instrument, specifically adaptive self-reflection, this study employed CFA (Confirmatory Factor Analysis). The initial first-order testing of thethree dimensions of the adaptive self-reflection measurement instrument, which originally consisted of 99 items, resulted in a model that was not fit in the initial analysis. Consequently, modifications were made to obtain an acceptable measurement model (unidimensional), reducing the number of fit items to 25. However, further testing of these 25 items still did not yield a fit measurement model, requiring additional modifications until a final unidimensional measurement model with 15 fit items was achieved. The limitation of this study is that CFA initially struggled to achieve a model fit, necessitating the removal of several items to ensure model acceptability. In contrast, employing Rasch offers greater flexibility in eliminating problematic items by evaluating infit and outfit values, providing a more precise and systematic approach to item refinement.

3.5 Wright Map

The validity of the construct can be assessed through the hierarchical arrangement of items, which is visually represented in a WrightMap. This map illustrates the distribution of item difficulty levels on the right panel and participants' abilities (latent traits) on the left panel. Items with lower difficulty are positioned at the bottom, while those of moderate difficulty are placed in the middle, marked by "M" on the right side. Conversely, items of higher difficulty are located at the top, as described by (Yu, 2020).

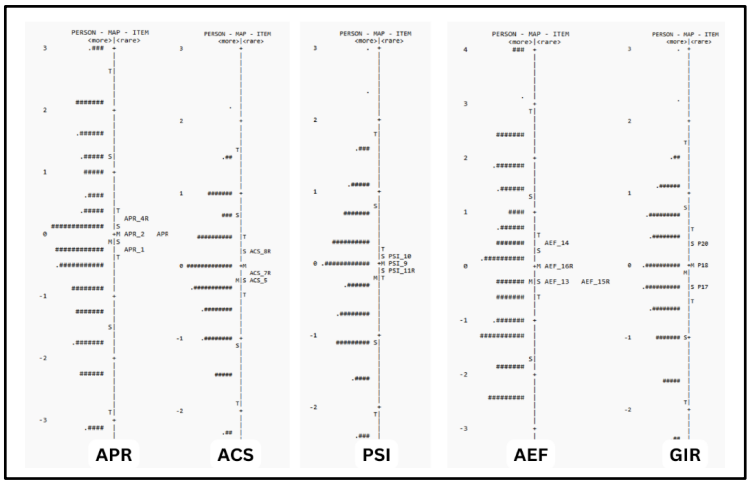

Figure 1. Wright Map

Referring to the Wright Map displayed in Figure 1, it is evident that the items across all five subscales are predominantly concentrated around a logit value of 0. This alignment suggests that the participants' abilities correspond well with the difficulty levels of the items, reflecting an appropriate balance between the test's demand and the participants' capabilities. However, the uneven distribution of item difficulty levels suggests limitations in the measurement tool's effectiveness. Specifically, while it is well-suited for accurately assessing participants with average to moderate abilities, it is less effective for individuals with low and high abilities. For such participants, the lack of adequately challenging items may lead to less precise measurements, highlighting the need for potential item adjustments to better accommodate a wider range of abilities and enhance the instrument's applicability.

A different result was observed in the study conducted by Bahiyah & Savitri, (2018). Although their research also focused on validating a measurement instrument related to individual self-adjustment, the CFA analysis used in this study only examined the relationships between latent variables without assessing whether certain items were too easy or too difficult for respondents. In contrast, Figure 1 provides a more in-depth insight into the difficulty level of each item, where items positioned higher on the scale are more difficult, while those lower are easier. Additionally, it illustrates how respondents with varying levels of ability responded to each item, with respondents positioned higher on the scale demonstrating greater ability, and vice versa.

3.6 DIF (Differential Item Functioning) by Gender

DIF reflects the tendency of participants to provide consistent responses to specific items under similar conditions (Hagquist & Andrich, 2017). This means that certain items may exhibit bias toward specific groups of participants (e.g., gender, age, or culture), causing differences in how these groups

interpret and answer the items correctly, consistently, and in alignment with the construct, even when individuals have the same ability level. DIF analysis can be conducted by examining the probability (p-value) score, if an item shows significant DIF, identified by a p-value less than 0.05 (sig. < 0.05), it suggests that the item is not fully impartial or neutral in assessing the intended ability. This lack of neutrality suggests potential item bias, affecting measurement accuracy and reliability across subgroups. Based on the DIF results, all items in this study had prob score (p > 0.05), indicating there is no DIF because there are no significant probabilities in either Welsc or Mantel in how male and female participants interpreted the items. The mixture rasch model analysis also reveals that from class 1 to class 2, there is no decrease in AIC and BIC across all sub-scales, indicating that no latent sub-classes exist. In contrast, a prior study by Jannah, (2013), which tested construct validity in a self-adjustment instrument employed the Confirmatory Factor Analysis (CFA) approach. However, CFA proved less effective in detectingsub-latent groups within a population, such as those with distinct response patterns, compared to the mixturerasch model.

3.7 Mixture Rasch Model

The Mixture Rasch Model analysis was conducted to uncover latent subclasses within the data, employing evaluation metrics like AIC, BIC and CAIC to ensure a thorough and reliable assessment. According to Nylund et al., (2007), BIC is often recommended as the primary criterion for determining the optimal number of latent classes, as it favors simpler models and reduces the risk of overfitting. Therefore, when adding sub-latentclasses, it is essential to evaluate the decrease in AIC and BIC, especially BIC due to its more conservative nature in model selection.

Tabel 6. Mixture Rasch Model

Sub-Scale	Class	AIC	BIC	CAIC	Log-likelihood
APR	1	1988	2062	2084	-972
	2	1989	2140	2185	-949
ACS	1	1721	1779	1796	-844
	2	1664	1782	1817	-797
PSI	1	1676	1733	1750	-821
	2	1622	1740	1775	-776
AEF	1	1852	1925	1947	-904
	2	1842	1992	2037	-876
GIR	1	1687	1745	1762	-827
	2	1642	1759	1794	-786

Tabel 7. Frequency of Genders

Genders	Membership	Count	Log-likelihood
Male	1	118	54.38%
	2	88	40.55%
Female	1	7	3.23%
	2	4	1.84%

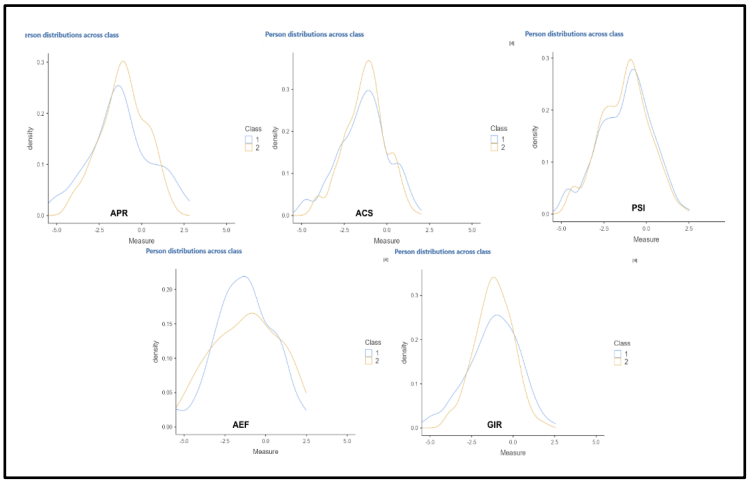

Figure 2. Person Distribution Across Class

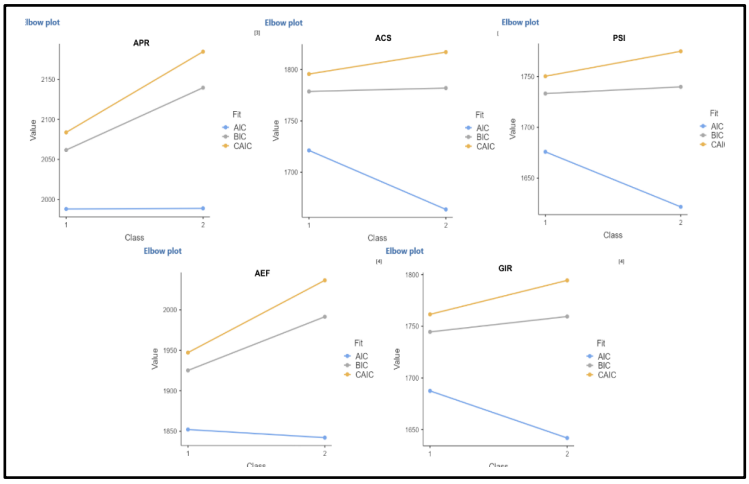

Figure 3. Elbow Plot

The analysis results indicate that from class 1 to class 2, there is an increase in AIC and BIC across all sub-scales, suggesting that no latent sub-classes exist. However, Morgan, (2015) emphasizes that researchers should avoid depending solely on one information criterion. Instead, they are encouraged to adopt a combination of multiple criteria while ensuring their findings are validated through robust theoretical frameworks and considerations. The Rasch Mixture analysis also revealed that males dominated all classes, with the highest proportion in class 1 (54.38%), while females had significantly smaller numbers across all classes, with the highest proportion (3.23%). Membership in the two classes showed an uneven distribution between males and females.

In comparison to the previous study by Faizah et al., (2020), which examined an adapted scale from the INCA (Inventory of New College Student Adjustment) by Watson & Lenz (2018), this study primarily focused on analyzing the relationship between self-esteem, resilience, and college adjustment among first-year students. However, a limitation of this study lies in the validation of the measurement instrument, as it did not include an evaluation of Item Difficulty or Differential Item Functioning (DIF). As a result, it remains unclear whether each item in the INCA scale functions differently across various groups, such as gender differences (male or female students) or variations in socio-economic background.

4. CONCLUSION

The construct validity analysis conducted utilizing the Mixture Rasch Model reveals that the findings not only provide insights into the extent to which the measurement tool aligns with the theoretical framework of self-adjustment but also include additional information such as the reliability of the self-adjustment measurement tool, demonstrating high internal consistency. The subsequent analysis indicates that the self-adjustment measurement used in this study is fit according to the Rasch Model without requiring any items to be replaced or removed based on the infit, outfit, point-measure correlation that meet the required criteria. The analysis of unidimensionality and rating scale yielded raw variance scores ranging from 40% to 62%, indicating that the measurement tool in this study tends toward unidimensionality. This suggests that participants are able to differentiate their response choices within the measurement tool. Another significant result related to item bias, as the analysisof DIF revealed that the majority of values does not display item bias, as male and female participants interpret the items similarly. The results of the Mixture Rasch Model analysis indicate that there is no observed decrease in AIC and BIC across all sub-scales, suggesting the absence of latent sub-classes within the data. Based on the research findings, the implications of the Self-Adjustment Questionnaire measurement can be utilized as a research assessment tool in both academic settings and institutional contexts. It can serve as a reference for interventions as well as for policy-making decisions.

This study has a limitation in the data collection process, which may not fully capture the characteristics of Indonesian university students due to the non-probability sampling technique. Another limitation highlights the uneven distribution of item difficulty and participant abilities, mainly due to the limited number of items, which hinders the instrument's capacity to distinguish between varying difficulty levels. Specifically, this limitation poses challenges in distinctly identifying variations among the most easiest and difficult items. As a result, the measurement tool may struggle to provide a comprehensive and

accurate assessment of participant abilities across the full spectrum, potentially affecting its effectiveness in capturing subtle differences in performance or traits. This highlights the need for an expanded and more balanced item pool to enhance the instrument's capacity for precise measurement. Even so, further research on the validation of self-adjstment measurement using the Mixture Rasch analysis shouldbe conducted by increasing the number ofparticipants and considering background characteristics such as gender, age, educational status, or diverse occupations. This would enable more accurate and representative measurement of the dimensions of the self-adjustment construct. Overall, the Self-adjustment Questionnaire can be effectively implemented in future studies and practical settings in both academic and institutional contexts with considerations, as explained before.

5. REFERENCES

Ababiel, M. W., Mahastuti, D., & Patrika, F. J. (2023). Penyesuaian diri, dukungan sosial teman sebaya dan student well-being pada mahasiswa vokasi pelayaran. *Jurnal Psikologi Poseidon, 6*(1), 66–79. https://doi.org/10.30649/jpp.v6i1.85.

Agestia, E., Safitri, D., & Sujarwo, S. (2024). Adaptasi Mahasiswa Dalam Mengatasi Culture Shock Dalam Perkuliahan. *RISOMA : Jurnal Riset Sosial Humaniora Dan Pendidikan, 2*(4), 253–264. https://doi.org/10.62383/risoma.v2i4.180.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

Araya, J. A. Z., Castro, V. S., Rojas, E. M., & Mora, T. E. M. (2018). Advantages of the Rasch Model for Analysis and Interpretation of Attitudes: the Case of the Benevolent Sexism Subscale. (2018). *Revista Evaluar, 18*(3), 1–13. https://doi.org/10.35670/1667-4545.v18.n3.22201.

Arifin, A. S. (2013). Studi kasus dampak penjurusan studi pilihan orang tua terhadap penyesuaian diri peserta didik SMAN 1 Kediri.

Avinç, E., & Doğan, F. (2024). Digital literacy scale: Validity and reliability study with the rasch model. *Education and Information Technologies, 29*, 22895–22941. https://doi.org/10.1007/s10639-024-12662-7.

Bahiyah, S., & Savitri, S. I. (2018). Validasi Struktur Internal Alat Ukur Refleksi Diri Adaptif melalui CFA. *JPsi: Indonesian Journal of Psychology, 45*(2), 107–131. https://doi.org/10.22146/jpsi.34966.

Bond, T. (2021). Applying the Rasch model: Fundamental measurement in the human sciences (4th ed.). Routledge.

Che Lah, N. H., Tasir, Z., & Jumaat, N. F. (2022). An Evaluation of the Online Social Learning Environment Instrument (OSLEI) Using Rasch Model Analysis. *SAGE Open, 12*(2). https://doi.org/10.1177/21582440221104083.

Credé, M., & Niehorster, S. (2012). Adjustment to College as Measured by the Student Adaptation to College Questionnaire: A Quantitative Review of its Structure and Relationships with Correlates and Consequences. *Educational Psychology Review, 24*(1), 133–165. DOI 10.1007/s10648-011-9184-5

Creswell, J. W., & Creswell, J. D. (2018). Research design: qualitative, quantitative, and mixed methods approaches (5th ed.). Thousand Oaks, CA: SAGE Publications.

Daulay, N. (2021). MOTIVASI DAN KEMANDIRIAN BELAJAR PADA MAHASISWA BARU. *Al-Hikmah: Jurnal Agama Dan Ilmu Pengetahuan, 18*(1), 21–35. https://doi.org/10.25299/al-hikmah:jaip.2021.vol18(1).5011.

Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi. (2024). Buku Panduan Penyusunan Kurikulum Pendidikan Tinggi Mendukung Merdeka Belajar-Kampus Merdeka Menuju Indonesia Emas.

Dwiliesanti, W., & Yudiarso, A. (2022). Rasch Analysis of The Indonesian Version of Individual Work Performance Questionnaire (IWPQ). *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia), 11*(2), 153-167. http://dx.doi.org/10.15408/jp3i.v11i2.24157.

Faizah., Marmer, F. V., Aulia, N. N., Rahma, U., & Dara, Y. P. (2020). Self-esteem dan resiliensi sebagai prediktor penyesuaian diri mahasiswa baru di Indonesia. *Persona: Jurnal Psikologi Indonesia, 9*(2), 335–352. https://doi.org/10.30996/persona.v9i2.3448.

Fauziana, A., & Wulansari, A. D. (2021). Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch. Jurnal Ibriez : Jurnal Kependidikan Dasar Islam Berbasis Sains, 6(1), 10–19. https://doi.org/10.21154/ibriez.v6i1.112.

Fisher, W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions, 21*(1), 1095.

Gluck, J., & Spiel, C. (1997). Item response models for repeated measures designs: Application and limitations of four different approaches. *Methods of Psychological Research, 2*(1), 1–19. https://doi.org/10.23668/psycharchives.12717.

Gunawan, W., & Yudiarso, A. (2023). Uji validitas konstruk work-life balance dengan rasch model. *Jurnal RAP (Riset Aktual Psikologi), 14*(1), 97–111. https://doi.org/10.24036/rapun.v14i1.120799.

Haber, A., & Runyon, R. P. (1984). Psychology of adjustment. Dorsey Press.

Handayani, Y., Rahmawati, R., & Widiasih, W. (2023). Using Rasch Model to Analyze Reliability and Validity of Concept Mastery Test on Electricity and Magnetism Topic. *JIPF: Jurnal Ilmu Pendidikan Fisika, 8*(2), 226-239. http://dx.doi.org/10.26737/jipf.v8i2.3877.

Hagquist, C., & Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and quality of life outcomes, 15*(1), 181. https://doi.org/10.1186/s12955-017-0755-0.

Holster, T. A., & Lake, J. (2016). Guessing and the rasch model. *Language Assessment Quarterly, 13*(2), 124–141. https://doi.org/10.1080/15434303.2016.1160096.

Jannah, M. (2013). Uji validitas konstruk pada instrumen penyesuaian diri haber & runyon (1964) dengan metode confirmatory factor analysis (cfa). *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia), 2*(6), 422–433. DOI: 10.15408/jp3i.v2i6.10752

Jiao, H., Lissitz, R.W., Macready, G., Wang, S., & Liang, S. (2011). Exploring levels of performance using the mixture Rasch model for standard setting. *Psychological Test and Assessment Modeling, 53*(4), 499–522.

Khine, M. S. (2020). Rasch measurement applications in quantitative educational research. *Springer Nature Singapore*, 47–73. https://doi.org/10.1007/978-981-15-1800-3_4.

Linacre, J. M. (1994). Sample Size and Item Calibration [or Person Measure] Stability.

Linacre, J. M. (2024). A user's guide to winstep® ministep rasch-model computer programs.

Linacre, J. M. (2002). What do Infit and Outfit, Mean-square and Standardized mean?.

Mamahit, H. C., & Deviani, P. (2022). Regulasi diri selama pembelajaran daring pada mahasiswa program studi bimbingan dan konseling angkatan 2018-2020 unika atma jaya jakarta. *Jurnal Psiko Edukasi, 20*(2), 120–132. https://doi.org/10.25170/psikoedukasi.v20i2.3922.

Morgan, G. B. (2015). Mixed mode latent class analysis: an examination of fit index performance for classification. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(1), 76–86. https://doi.org/10.1080/10705511.2014.935751.

Muntazhimah, M., Putri, S., & Khusna, H. (2020). Rasch Model untuk Memvalidasi Instrumen Resiliensi Matematis Mahasiswa Calon Guru Matematika. *JKPM (Jurnal Kajian Pendidikan Matematika), 6*(1), 65–74.

McCreary, L. L., Conrad, K. M., Conrad, K. J., Scott, C. K., Funk, R. R., & Dennis, M. L. (2013). Using the rasch measurement model in psychometric analysis of the family effectiveness measure. *Nursing research, 62*(3), 149–159. DOI: 10.1097/NNR.0b013e31828eafe6

Nur, L., Yulianto, A. Suryana, D., Malik, A. A., Ardha, M. A. A., & Hong, F. (2022). An analysis of the distribution map of physical education learning motivation through Rasch modeling in elementary school. *International Journal of Instruction, 15*(2), 815–830. https://doi.org/10.29333/iji.2022.15244a.

Nyimas, H. A. S., & Rulanggi, R. (2022). Analisis Student Adaptation to College Qestionnaire (SACQ) sebagai Instrumen Pengukuran Penyesuaian di Perguruan Tinggi pada Mahasiswa Baru. *Buletin Poltanesa, 23*(1), 112–117. https://doi.org/10.51967/tanesa.v23i1.1247.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535–569. https://doi.org/10.1080/10705510701575396.

Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2019). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth, 25*(1), 104–112. https://doi.org/10.1080/02673843.2019.1596823.

Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema, 26*(1), 108–116. https://doi.org/10.7334/psicothema2013.260.

Rost, J., Carstensen, C., & von Davier, M. (1997). Applying the mixed rasch model to personality questionnaires. In j. Rost & r. Langeheine (eds.), applications of latent trait and latent class models in the social sciences. Münster: Waxmann.

Rost, J. (1990). Rasch models in latent classes: an integration of two approaches to item analysis. *Applied Psychological Measurement, 14*(3), 271–282. https://doi.org/10.1177/014662169001400 30.

Saalino, U., Razak, A. ., & Nur Hidayat Nurdin, M. . (2022). Hubungan Antara Sense Of Community Dan Penyesuaian Diri Pada Mahasiswa Baru Di Papua Barat. *Jurnal Psikologi Teori Dan Terapan, 13*(2), 155–163. https://doi.org/10.26740/jptt.v13n2.p155-163.

Safitri, R. M. (2020). Rasch Model Application to examine the Psychometric Properties of Premarital Sexual Behaviour Scalogram. *Journal of Psychology and Instruction, 4*(2), 52–59. https://doi.org/10.23887/jpai.v4i2.24941.

Sasti, K. I., Rismansyah., & Oktariansyah. (2023). Pengaruh efikasi diri dan penyesuaian diri terhadap kepuasan kerja pegawai pada pt. karya mandiri propertindo utama. *Jurnal Media Wahana Ekonomika, 20*(3), 485–496. https://doi.org/10.31851/jmwe.v20i3.13444.

Setiawati, F. A., Amelia, R. N., Sumintono, B., & Purwanta, E. (2023). Study item parameters of classical and modern theory of differential aptitude test: Is it comparable?. *European Journal of Educational Research, 12*(2), 1097–1107. https://doi.org/10.12973/eu-jer.12.2.1097.

Setiyawan, A. (2020). Faktor-faktor yang mempengaruhi reliabilitas tes. *AN NUR: Jurnal Studi Islam, 6*(2).

Smith, E. V., Jr., Conrad, K. M., Chang, K., & Piazza, J. (2002). An introduction to rasch measurement for scale development and person assessment. *Journal of Nursing Measurement, 10*(3), 189–206. DOI: 10.1891/jnum.10.3.189.52562

Sumintono, B. (2016). Penilaian keterampilan berpikir tingkat tinggi : aplikasi pemodelan rasch pada asesmen pendidikan. *Seminar Nasional Pendidikan IPA, FKIP Jurusan PMIPA*, 1–19.

Sumintono, B., & Widhiarso, W. (2015). Aplikasi pemodelan rasch pada assessment pendidikan. Trim Komunikata.

Syifa, S. A. M., Sartika, D., & Saefudin, M. A. (2024). Pengaruh Self-Efficacy terhadap College Adjustment pada Mahasiswa Baru Rantau di Kota Bandung. *Bandung Conference Series: Psychology Science, 4*(2), 061–1067. https://doi.org/10.29313/bcsps.v4i2.12962.

Tanujaya, L., & Yudiarso, A. (2023). Validity test of the indonesian version of the multicultural attitude scale (masque) questionnaire using the rasch model. *Humanitas: Jurnal Psikologi, 7*(2), 139–150. https://doi.org/10.28932/humanitas.v7i2.7420.

Toker, T., & Green, K. (2021). A comparison of latent class analysis and the mixture rasch model using 8th grade mathematics data in the fourth international mathematics and science study (timss-2011). *International Journal of Assessment Tools in Education, 8*(4), 959–974. https://doi.org/10.21449/ijate.102421.

Toker, T. & Seidel, K. (2023). A mixture rasch model analysis of data from a survey of novice teacher core competencies. *International Journal of Contemporary Educational Research, 10*(1), 147-156. https://doi.org/10.52380/ijcer.2023.10.1.349.

Vermunt, J. K. (2022). Latent class analysis. In P. Peterson, E. Baker, B. McGaw, F. Rizvi, G. Smith, & K. Gutierrez (Eds.). International encyclopedia of education (4th ed.). Oxford: Elsevier.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars & A. McCutcheon (Eds.), Applied latent class analysis. Cambridge University Press.

Wahyuni, A. (2022). DETECTION OF GENDER BIASED USING DIF (DIFFERENTIAL ITEM FUNCTIONING) ANALYSIS ON ITEM TEST OF SCHOOL EXAMINATION YOGYAKARTA. *Jurnal Evaluasi Pendidikan, 13*(1), 46–49. https://doi.org/10.21009/jep.v13i1.26554.

Wang, C. Xu. G., & Shang, Z., & Kuncel, N. (2018). Detecting aberrant behavior and item preknowledge: a comparison of mixture modeling method and residual method. *Journal of Educational and Behavioral Statistics, 43*(4), 469–501. https://doi.org/10.3102/1076998618767123.

Widhiarso, W. (2010). Model rasch campuran dalam teori respons butir. *Ids Working Papers*, 1–4.

Wistarini, N. N. I. P., & Marheni, A. (2019). Peran dukungan sosial keluarga dan efikasi diri terhadap stres akademik mahasiswa baru Fakultas Kedokteran Universitas Udayana angkatan 2018. *Jurnal Psikologi Udayana, 6*, 164–173.

Yu, C. H. (2020). Objective measurement: how rasch modeling can simplify and enhance your assessment. *Rasch Measurement*, 47–73. https://doi.org/10.1007/978-981-15-1800-3_4.

Yulisharyasti, L., Nurdin, A., Aulia, N., Arfa, F.A., & Fadjryani. (2023). Analyzing the quality of measurement instruments of multiple choice questions on class xi economics material in public high school 3 gorontalo through classical test theory and rasch models. *Parameter: Journal of Statistics, 3*(1), 28–38. https://doi.org/10.22487/27765660.2023.v3.i1.16417.

Zhang, J., & Goodson, P. (2011). Predictors of international students' psychosocial adjustment to life in the united states: a systematic review. *International Journal of Intercultural Relations, 35*(2), 139–162. https://doi.org/10.1016/j.ijintrel.2010.11.011.