International Journal of Psychology and Educational Studies, 2025, 12(3), 169-179

IJPES

www.ijpes.com

International Journal of Psychology and Educational Studies

ISSN: 2148-9378

The Contribution of Rasch Modeling on Measuring Attitudes For Better Classroom Assessment

Akif AVCU[1]

[1]*Faculty of Education, Marmara University, İstanbul, Türkiye* 0000-0003-1977-7592

ARTICLE INFO

Article History
Received 07.08.2024
Received in revised form 05.09.2024
Accepted Click
Article Type: 30.11.2024

ABSTRACT

This-review explores the significant contributions of Rasch modeling in enhancing classroom assessment practices, particularly in measuring student attitudes. Classroom assessment has evolved from standardized testing to integrative practices that emphasize both academic and affective dimensions of student development. Accurate attitude measurement tools are crucial for understanding students' interests, motivations, and dispositions, which influence their learning outcomes and overall academic engagement.Rasch modeling, a robust framework within Item Response Theory, addresses limitations of traditional Likert-scale analysis by converting ordinal data into interval-level measurements. It enables precise calibration of measurement tools while ensuring unidimensionality, reliability, and the elimination of item bias. Studies reviewed demonstrate the effectiveness of Rasch modeling in diverse educational contexts, from primary school coding attitudes to secondary science engagement and university-level course perceptions.This review highlights the utility of Rasch models in developing psychometrically sound scales, refining item properties, and diagnosing bias in assessments. It underscores the transformative potential of Rasch analysis in classroom assessments, offering educators actionable insights for creating reliable, valid tools tailored to various learning environments. The findings advocate for wider adoption of Rasch modeling in educational research to enhance both theoretical understanding and practical applications of student assessment.

Keywords:
Classroom assessment; Rasch modeling; attitudes

1. Introduction

The transformation of educational assessment from standardized testing to classroom based activities

The activities that a teacher performs during the day are both multidimensional and varied. When we consider them dimensionally, the legal regulations to be followed and the routine operations within the school come to the fore. Especially when we look at the activities within the school, teachers have many duties such as carrying out teaching activities, accompanying students' social and cultural activities, maintaining discipline within the school and being a behavioral model for students. In addition to these, another task of teachers is to evaluate students' achievements in exams, in class and in terms of various performances. Each of these activities is valuable individually. The focus of this study is on the assessment of students. Accurate assessment is first and foremost a part of healthy communication and a prerequisite for effective feedback (Mertler, 2016). As a matter of fact, it is very important for students to receive regular feedback about their products, exams and performances in the process of developing themselves. In addition, this feedback that they can receive with regularity will support them to be more intrinsically motivated as a stimulus (Ryan, Connell, & Deci, 1985). On the contrary, when students receive inaccurate feedback, it affects their emotional state, attitudes and enthusiasm towards the course, school and teachers (Brookhart & McMillan, 2020).

[1]Corresponding author's address: Marmara University, Faculty of Education, İstanbul, Türkiye
e-mail: avcuakif@gmail.com
Citation: Avcu, A. (2025). The contribution of Rasch Modeling on measuring attitudes for better classroom assessment. *International Journal of Psychology and Educational Studies*, *12*(3), 169-179. https://dx.doi.org/10.52380/ijpes.2025.12.2.1432

The process of student assessment involves, first of all, gathering information from many different sources, which is shaped by the expected goals and objectives of the lesson or activity. This information is then synthesized and interpreted by teachers and used to make decisions Airasian (2000). This decision can be in the form of revision of classroom practices or it can be used to determine the student's transition to the next level. The information gathering process of this process can be carried out through standardized tests or through a variety of classroom assessment activities. In general terms, classroom assessment is used to refer to the activities carried out to determine the objectives that are intended to be achieved within the scope of the course and the extent to which these objectives are met. These activities are very important to support the decisions to be taken as part of a large assessment system (Mertler, 2016). The accuracy of these decisions depends on the data coming from different sources. Soland (2013) stated that an effective assessment system depends on the information coming from different sources at the decision-making stage. Therefore, teachers should not use data from a single assessment source to make judgments about a student's learning, attitudes or dispositions. Judgments should always be made by considering multiple sources.

Indeed, Stiggins & Bridgeford (1985) described the importance of informal observation, which often takes place spontaneously, information obtained from verbal questions asked during class lectures and observations of classroom activities, in the decision-making process about students. This approach is also based on the view that assessment and instruction are inseparably part of a holistic process (Graue, 1993),

This holistic understanding of instructional activities and the assessment process as inseparable has also changed the traditional understanding of the concept of assessment. Especially with the increasing adoption of the constructivist learning approach, classroom assessment has been broadened in terms of its role in instructional processes rather than being summative learning (Penuel & Shepard, 2016). Accordingly, the assessment process has become integrated with the instructional processes in the classroom and transformed into a supportive activity to help students learn and improve teaching. With this transformation, the concept of classroom assessment has come to the fore.

Classroom assessment is an approach that focuses on student development, such as identifying students' strengths and weaknesses, taking the necessary measures to compensate for deficiencies, and providing feedback to both students and parents while teaching activities continue (McMillan, 2013). Classroom assessment methods can be very formal or very informal. Whether the assessment is formal or informal is related to the spontaneous/random nature of the assessment (Oosterhof, 1999). Formal assessment methods are planned before the assessment is carried out. This is not the case with informal assessment.

Although classroom assessment has been discussed so far, generally implying the level of knowledge in academic subjects, in fact students' affect is also an important part of their academic and social emotional development. It is important to identify these other factors involved in the learning process because only in this way it is possible to design an appropriate curriculum and improve students' academic outcomes (Zeegers, 2004). One of the main determinants of the teaching process is students' attitudes towards teachers, school, teaching methods and course content (Green et al. 2012). Affective factors are as important as ability for the effectiveness of education (Steiner & Sullivan, 1984). Pell and Jarvis (2001) stated that attitudes have an impact on students' attention, consistency and behavior in the classroom. For example, Mao et al. (2021) found a statistically significant and strong positive relationship between students' attitudes towards science and their academic achievement in science. In other words, by determining student attitudes, it is possible to predict students' declining interest in a course and to take some measures in this direction. In other words, there is a need for valid measurement tools to determine students' attitudes for all kinds of courses, in-class activities and even instructional models that involve instructional processes.

Moreover, in classroom situations, teachers need to make sure that tests or other assessment tools are able to measure students' attitudes effectively in the assessment process. Therefore, an important part of the classroom assessment process is the monitoring of students' affective characteristics. For these reasons, classroom activities are nowadays designed to include attitudes and various techniques have been proposed (Angelo & Cross, 2012). One of the methods of measuring attitudes is the use of measurement tools.

Rasch Modeling: a brief introduction

As with any measurement tool used in education and psychology, the measurement tools of attitude surveys that can be used in classroom assessment need to be valid and reliable. Traditionally, factor analysis is performed to assess the validity of a measurement instrument (Mešić et al., 2019). gTo assess reliability, Cronbach alpha internal consistency coefficient and test-retest are used to assess the stability of the measurement. On the other hand, it is known that modern approaches such as Item Response Theory are widely used in the process of measurement instrument development/revision. Since it is beyond the scope of this article to provide comprehensive theoretical information about Item Response Theory, interested readers are recommended to review Reid et al. (2007). Rasch models, one of the important models within the scope of item response theory, come to the forefront in the process of scale development and adaptation.

As it is known, in most attitude surveys, students are presented with a set of statements as a stimulus and asked to express their opinions about the statements. In order to express their opinions, they are usually asked to mark the most appropriate option on a Likert scale ranging from strongly disagree to strongly agree. However, when the scales are developed, the scale is considered at the interval level in the analysis phase. But in fact Likert scales are ordinal. Therefore, this assumption is not acceptable in the process of calculating total scores. In this process, it is inevitable that the statistical analysis will give erroneous results. For this reason, the use of Rasch models has become widespread because Rasch models make it possible to convert ordinal scales into interval-level data. The Rasch model (Rasch, 1960) is one of the most recognized approaches within item response theory (IRT). It provides a method for calibrating ordinal data onto a common measurement scale while testing key assumptions such as unidimensionality, linearity, and local independence (Wright, 1996). Unlike factor analysis, which focuses on reducing data into a few latent factors, the Rasch model emphasizes the detailed interaction between individuals and items (Reckase, 1997). This method makes it possible to convert ordinal data into interval-level measurements, allowing for more nuanced insights into item behavior across diverse respondents and contributing to the overall reliability of the resulting scale. In the context of the Rasch model, different formulations are available to estimate response probabilities, and these models vary depending on the structure of the item characteristic function. This function serves as a model for understanding the relationship between an individual's item responses and their underlying ability level (Bond & Fox, 2015).

The Rasch model also enables researchers to assess reliability for both at person level and at item level. The item reliability index indicates how stable item difficulty estimates remain when the same items are administered to a comparable population. The item discrimination index reflects how well the items differentiate along the latent trait being measured. Likewise, the person reliability index gauges the consistency of individuals' standings on a parallel form of the assessment. Finally, the person discrimination reliability provides insight into how effectively respondents can be distinguished based on the construct being evaluated.

At the same time, another output that can be obtained from Rasch modeling is the Wright map. This map is a tool used to predict the difficulty levels of test items and to express a student's performance based on these test item difficulty levels. The Wright map allows researchers to assess how well test items describe a variable. The Wright map help researchers in comparing the predicted order of item difficulty with the real order of item difficulty and thus assess the theoretical appropriateness of the measurement.

Aim of the study

The aim of this review is to review the studies in which Rasch analysis was used while developing, revising or evaluating the psychometric properties of measurement tools used in measuring attitudes, which have an important place for classroom assessment practices. It also aims to draw the attention of practitioners who will be developing measurement tools for classroom assessment practices and to provide them with an outline for this process. the theoretical appropriateness of the measurement.

2. Methodology

This study was conducted as a review because it includes an overview on the use of Rasch modeling for scale development/revision purpose by providing theoretical knowledge and recent empiricalstudies. Considering the inclusiveness of many different databases, all article searches were conducted through Google Scholar search engine. While conducting the search, the criteria of developing/revising the scale with the Rasch model or examining the psychometric properties were taken into consideration. During the search, some concepts such as course interest, which are synonyms/close synonyms of attitude, were not taken into consideration and the search process continued by considering only the keyword attitude. Since attitude is a general concept in the instructional context, there was no restriction in this direction as there could be attitudes towards the school, attitudes towards the course or attitudes towards a specific subject within the course. At the same time, although classroom assessment is generally used more intensively in the early years of education, it is also used at higher education level, so no filtering based on educational levels was made. The articles that used only other parametric IRT models or that don't use Rasch models for an aim other that scale development/revision were excluded from the study. Since the findings are summarized and evaluated in accordance with the flow of this study, some of the articles found in the literature are included in a way to preserve diversity. All in all, a total of thirteen article was included in the current review. The şhoudl note that, only a smally number of articles were included due to the redundance of articles in the literature. The table 1 include the summary of articles.

3. Findings

In this part, the studies included in ths current research were summarized and discurred in terms of the Rasch outputs they provide and contribution these outputs to the readers/test users. The detailed summary of these studies can be find in Table 1

The first study (Alasgarova, 2022) examined the psychometric properties of the "Primary School Students' Coding Attitudes Questionnaire" developed to assess primary school students' attitudes towards coding using Rasch modeling. This scale consists of 23 items. In addition, the scale consists of 5 different sub-dimensions of attitudes towards coding. In this study, each dimension was analyzed separately. It was concluded that the construct validity of the scale was at a sufficient level, especially by evaluating the item fit in the perspective of Rasch analysis. No incompatibility was found for any of the scale items. Therefore, it was concluded that this scale has high psychometric properties in assessing attitudes towards coding. Finally, considering that data obtained from multinational assessment exams are used extensively in the relevant literature, it can be said that this study is valuable in terms of focusing on a specific learning area.

In another study conducted with secondary school students, Suryadi et al. (2023) sought evidence for the validity of the Attitudes Towards Science questionnaire used to assess the attitudes towards science of secondary school students studying in Indonesia by examining its psychometric properties. In addition, differential item functioning was examined to assess whether there were any items that caused bias in terms of gender and academic level of the students. Although the scale consisted of four subscales, analyses were conducted to assess whether it was unidimensional. The final sample group consisted of 223 students. As a result of the analyses based on the Rasch model, it was concluded that the attitude towards science instrument was not unidimensional and it was understood that two dimensions could be combined. As a result, the existence of a three-dimensional structure was finally confirmed. In addition, it was determined that the reliability of the sub-dimensions was quite high, but it was concluded that the reliability of the measurement tool was relatively lower in the dimensions of expectation and perception towards scientists. In other words, the attitude towards science scale provides less information about individuals with positive and negative perceptions of scientists and therefore fails to differentiate students based on their attitudes. In addition, the researchers concluded that two items of the scale showed gender bias and made suggestions to reduce the items and warned the practitioners at this point. Considering that attitude towards science is an important affective factor that can determine students' interest and success especially in science courses, it is thought that this study can make important contributions in terms of making more precise and reliable measurements of attitude towards science.

Table 1. *Summary of the Articles Included in the Study*

Authors	Scales Used	Types of Rasch Analysis Outputs
Alasgarova, G. A. (2022)	Student Coding Attitudes Survey (SCAS)	Item difficulty, person ability estimates, item fit, reliability indices, category thresholds
Suryadi, A., Fawaiz, S., Kurniati, E., & Swandi, A. (2023)	Modified Attitudes Toward Science Scale (MATS)	Differential item functioning (DIF), item-person map, item fit statistics, reliability indices
Siegel, M. A., & Ranney, M. A. (2003)	Changes in Attitude About the Relevance of Science (CARS)	Item difficulty, reliability indices, person-item maps, fit statistics
Huang, L., Huang, F., Oon, P. T., & Mak, M. C. K. (2019)	Student Attitudes Toward Science Scale	Dimensionality analysis, item difficulty, person reliability, fit statistics
Muhammad, I., et al. (2023)	Student Interest in Mathematics Learning (SIML) Scale	Item-person map, reliability indices, item fit statistics, threshold analysis
Sukri, A., et al. (2022)	Green Character Instrument	Factor analysis integration, item difficulty, reliability indices, category thresholds
Oon, P. T., & Fan, X. (2017)	Science Attitude Rating Scales	Psychometric improvement analysis, dimensionality, item difficulty, reliability indices, fit statistics
Sabah, S., Hammouri, H., & Akour, M. (2013)	Attitudes Toward Science Scale	Cross-country DIF, item fit statistics, person-item maps, reliability indices
Pey Tee, O., & Subramaniam, R. (2018)	TIMSS 2011 Attitudinal Data	DIF analysis across countries, item fit, person reliability, scale psychometric properties
Zain, A. N. M., et al. (2010)	Students' Attitudes Toward Science Scale	Item difficulty, reliability indices, item fit, person-item maps
Akour, M. M. (2022)	Survey of Attitudes Toward Statistics	Rating scale analysis, item difficulty, reliability indices, fit statistics
Papanastasiou, E. C., & Schumacker, R. (2014)	Attitudes Toward Research Scale	Rating scale analysis, dimensionality, item fit statistics, reliability indices
Veas Iniesta, A., et al. (2023)	Spanish adaptation of School Attitude Assessment Survey-Revised	Multidimensional Rasch analysis, item fit, person-item maps, reliability indices, DIF analysis

In another study (Siegel & Ranney, 2003), which measured students' attitudes towards science and used Rasch modeling, a new scale was developed rather than evaluating an existing scale and Rasch modeling was used for this purpose. In addition, since this study aimed to examine the effects of different curricula on students' attitudes, equivalent forms were also developed based on the findings obtained from Rasch modeling. From this perspective, it can be said that Rasch modeling has the potential to produce equivalent forms that can be used in classroom assessment processes. In this way, in this study, the extent to which innovative teaching activities supported students' science attitudes using traditional and Rasch modeling techniques was evaluated.

Another study (Huang et al., 2019) includes a total of 1,133 middle and high school students from China were reached in a study based on the Asian Students' Attitudes Towards Science Course. In this study, the data were examined in terms of compliance with the Rasch model, which requires invariant and consistent response category functioning.

In a study conducted by Muhammad et al. (2023), it was aimed to collect information about mathematics teaching students' interest in learning mathematics by using an augmented reality environment. In this study, 20 mathematics teaching students identified by purposive sampling method were included. The data were collected using a student interest inventory consisting of 18 statement items analyzed using the Rasch model. The high logit value obtained at the end of the study was interpreted as high student interest in mathematics. As can be seen, Rasch analysis was used in this study to determine item difficulty levels.

In addition, Sukri and his collegues (2022) aims to develop and validate a green character measurement tool that aims to determine students' environmental behavior and awareness. This study is rich in that it includes different statistics and analyses compared to other studies using Rasch analysis and shows that the analyses based on the Rasch model are inferentially rich. This scale comprises 40 statement items divided into five dimensions: private pro-environmental behavior, public pro-environmental behavior, environmental knowledge, environmental values, and environmental attitudes. The dataset includes responses from 1,398 students from various universities across Indonesia. In the process of examining the construct validity, the results of explanatory and confirmatory factor analyses were also included in addition to the use of the Rasch model. As a result of these analyses aimed at examining construct validity, six items were eliminated and thirty-four items belonging to five sub-dimensions were included in the final measurement tool. The researchers defined the final measurement tool as being easily understood by the participants and having a good item difficulty level. In addition, the assessment of item bias (Iseppi et al., 2021) shows that the items in this study were able to measure the green character of both male and female students. In addition, reliability levels were examined within the scope of the study and evidence was obtained to prove that the scale was reliable. Furthermore, in addition, Wright's map analysis was conducted to determine the difficulty level of the items (Scoulas et al., 2021), followed by rating scale analysis to assess the clarity and ease of interpretation of the response set in the instrument (Kim & Kyllonen, 2006), and results were obtained to support the psychometric properties of the scale, which were considered adequate.

Students' attitudes towards science is a prominent research area in science education literature. At the same time, there has been a growing interest in examining the measurement tools developed for science from a Rasch perspective. Many studies of attitudes towards science have used alternative expressions such as confidence, enjoyment or importance of science to examine this construct. Since the relationship between these statements and attitude is based on subjective judgment, these types of studies were not considered in the present study.

When the related literature is examined, it is seen that there are studies conducted using multinational data. Especially in large-scale international assessments such as TIMSS (Trends in International Mathematics and Science Study) and PISA (Program for International Student Assessment), the inclusion of items to determine students' science attitude levels has been a determinant of the preferences in the studies. However, it was observed that the number of studies conducted in the related literature was not sufficient. However, it is clear that these studies provide statistical advantages since it is possible to work with large sample groups.

Especially in recent years, science education researchers have developed a wide variety of attitudinal instruments, conducted validity and reliability studies, and these studies are used today. Within the scope of these studies, many data were collected with different statistical techniques, the data were analyzed and the findings were published. The results based on these publications have ultimately influenced public policies, educational practices, and the opinions and actions of many other stakeholders. Unfortunately, most science education tools have been developed using a theoretical measurement framework as a guide.

The first other study to be mentioned from studies derived from multinational assessment programs is the study conducted by Oon and Fan (2017). This study sought to illustrate the use of Rasch analysis in obtaining psychometric insights into rating scales designed to measure attitudes toward science. The analyses were conducted using a ready-made dataset collected as part of the Trends in International Mathematics and Science Study (TIMSS) (2011), which aims to assess students' attitudes towards science. This data set consists of a 20-item scale of attitudes towards science. The participants of the study were all eighth grade students from Hong Kong and Singapore (N = 9942). As a result of these analyses, features such as the invariance measurement of the attitude scale, the unidimensionality of its structure towards science, the optimal use of rating categories, and the item difficulty hierarchy in the scale were examined, and in light of these findings, suggestions were made on how the items of the attitude scale could be better designed. As can be seen, it is not possible to obtain these analyses with the traditional psychometric approach, and Rasch analysis has provided information that contributes to the measurement of attitude towards science with the advantages it provides.

In another similar but richer study (Sabah et al. 2013), a validity study was conducted based on TIMSS 2007 dataThe researchers aimed to offer insights and recommendations for other scholars studying attitudes toward science. They applied the Rasch measurement model to evaluate the scale's validity and reliability and to compare its performance across different countries. To ensure diverse data representation, the study included data from three top-performing countries (Singapore, Taiwan, and Japan), three low-performing countries (Qatar, Botswana, and Ghana), and three mid-performing countries (Australia, Scotland, and Italy). To assess the scale's construct validity, various statistics were calculated using the Rasch model. It was observed that the majority of the items fit the Rasch model. They also used these statistics to eliminate the items that did not fit. They also repeated the analysis with the remaining items. This study is valuable in showing that Rasch analysis can be used for validation and item elimination of attitude scales. Also, the comparison of successful and unsuccessful countries is meaningful to see how the psychometric properties of the scale may change at different levels.

In another study (Pey Tee & Subramaniam, 2018), secondary school students' attitudes towards science were conducted as a comparison study involving three countries: England, Singapore and the U.S.A. The researchers stated the reason for choosing these countries as a comparison of east and west. The data sets obtained from TIMSS 2011, which were adjusted for missing data, were used. In the study covering these three countries, data of 20,246 students were finally included in the study. Rasch model was used to analyze the data. Within the scope of the studies, evidence was sought for the validity and reliability of the scale of attitudes towards science and a comparison of the countries was made in certain aspects. Differential item functioning analyses revealed that some negative items may cause bias and some suggestions were made to the TIMMS developers in this direction. Furthermore, evidence for the unidimensionality of the measurement tool was obtained and its reliability level was reported to be adequate.

In another study (Zain et al., 2010) conducted in Malaysia to determine students' attitudes towards science learning in different secondary schools with a predominance of low-achieving students, the researchers reached 214 students and collected data using a survey. Using Rasch analysis, the findings of the study showed that students in low-performing schools believed that science was useful for society and found learning science interesting. On the other hand, in contrast, they had less positive self-concept about science and had insufficient awareness of science as a field that would benefit their future careers. Based on these findings, suggestions on how science teaching should be designed were presented.

Although relatively fewer, there are studies examining the psychometric properties of scales developed for this population by taking into account the negative attitudes of university students towards courses. In the first study (Akour, 2022), a questionnaire developed to measure university students' attitudes towards statistics was examined using Rasch modeling. The researchers stated that this questionnaire is widely used and that they conducted their study to accurately determine attitudes towards statistics. The study included 423 university students. The data collected within the scope of the research were collected by convenient sampling method. Using the Rasch rating scale model, this study aimed to determine the dimensionality, the fit of the items to the Rasch model, the item and person reliabilities, the functionality of the response categories, and how the items of this scale measuring students' attitudes towards statistics are distributed along the continuum of attitudes towards statistics. The findings showed that the items of the scale had excellent item agreement. At the same time, the researcher examined the interrater reliability and again found results that supported the high psychometric properties of the instrument. At the same time, Rasch analysis confirmed that the scale showed a unidimensional structure. In addition, it was found that the response categories functioned well. On the other hand, it was suggested that some items in some sub-dimensions could be changed in terms of content and some items could be completely removed from the scale. In addition, it was observed that the items were not well distributed across the ability spectrum and some suggestions were made to add new items to the measurement tool to make this distribution more homogeneous. Finally, in order to have a more effective rating scale, it was suggested that the number of response categories should be reduced to five instead of seven.

As a matter of fact, in a study (Papanastasiou & Schumacker, 2014), university students had negative attitudes towards research methods courses and for this reason, data were collected using a measurement tool that measures attitudes towards research. Afterwards, the data were analyzed and the psychometric properties of the scale were examined. In particular, it is important to include the variable of attitude towards research because attitude towards research is critical to university students' success and learning in many different contexts and courses. In addition, this study can provide valuable information for academics who provide research methods courses and more effective learning can be realized. Finally, the researchers concluded that this measurement tool, which consists of thirty items, is valid and reliable and that there is a need to increase students' attitudes.

Although classroom assessment is a process that takes into account educational activities, when we consider this assessment process in a contextual way, it can be said that another determinant general attitude that includes educational activities is the attitude towards the school. Ultimately, if the student develops a negative attitude towards the school in which he/she is enrolled, the function of classroom assessment activities at the course or subject level will be limited. In this direction, this section finally includes a study (Veas Iniesta et al., 2017) that examines the psychometric properties of a measurement tool developed to measure attitude towards school based on Rasch modeling. The revised version of the School Attitude Evaluation Questionnaire was used in the study. The researchers stated that the purpose of this study was to verify the dimensionality of the scale using multidimensional Rasch modeling. The data used in the study were obtained from 1398 students attending different high schools. The difficulty parameter and the ability parameter based on the data was obtained after calibration. As a result of the analyses, the model fit levels of ten items were found to be insufficient and they were removed from the scale. In addition, the differential item functioning analysis showed that the scale did not reveal a gender bias. Finally, it was concluded that the scale is reliable with a 7-category structure, compatible with the Rasch model, and has measurement precision. This study differs from other studies in the related literature in terms of using the multidimensional Rasch model and showing how the qualities of the scale can be improved with such a modeling.

4. Discussion, Conclusion and Recommendations

The literature review shows that the psychometric properties of many different attitude scales used in instructional processes have been examined using Rasch models. It is seen that parameter estimation, which is the first step of Rasch analysis, was performed in all of the measurement tools. Considering these parameters, the level of "difficulty" of the measurement tool was determined. Furthermore, in one article (Akour, 2022), it was interpreted whether the items homogeneously represent the ability spectrum. No such inference was found in other articles. Particular attention should be paid to this issue in scale development studies conducted to determine the final items.

In addition, in articles, item removal was generally recommended (Akour, 2022) or realized (Veas Iniesta et al., 2017) based on item agreement statistics. Although there is a statistical framework in classical test theory when deciding to remove items, the Rasch model, which is a model and probability-based approach, offers a broader decision framework when considered together with the other statistics it provides at this point.

In addition, some studies quantified differential item functioning at the item level (Suraydi et al., 2023; Pey Tee & Subramaniam, 2018; Veas Iniesta et al., 2017). In all of these studies, differential item functioning was examined in terms of gender variable and bias was evaluated according to this variable. On the other hand, only one study (Pey Tee & Subramaniam, 2018) observed a problem due to gender bias. Although it is not necessary to use Rasch models to test differential item functioning, the availability of such a possibility in the Rasch family of analyses enriches the analyses conducted in this direction.

Moreover, studies also vary in terms of sample sizes. Especially studies conducted with data from multinational benchmark tests can reach large data sets by taking advantage of the large sample sizes in these benchmark programs. On the other hand, since Rasch analyses are the simplest of the probabilistic analyses, large samples are not needed. In the articles reviewed, it was seen that Rasch modeling can be performed with a sample size of 200. This is an achievable number for many practical situations.

In conclusion, even though the traditional objective testing movement and modern classroom assessment techniques seem to be opposites, classroom assessment practices should use objective attitude tests to some extent. There are many different statistical models for the development of objective tests in the modern era. These models can provide solutions that allow for unbiased measurement of high precision with far fewer items. There are also many more sophisticated models based on item response. For this reason, there is a need for future studies to examine these models and to compile their benefits in the development and revision of attitude scales.

5. References

Airasian, P. W. (1989).Classroom assessmentanducational improvement. In L. W. Anderson (Ed.), *The effective teacher* (pp. 333-342). Random House.

Airasian, P. W., & Abrams, L. M. (2000). The theory and practice of portfolio and performance assessment. *Journal of Teacher Education*, *51*(5), 398-402. https://doi.org/10.1177/0022487100051005007

Akour, M. M. (2022). Rasch rating scale analysis of the survey of attitudes toward statistics. EURASIA Journal of Mathematics, *Science and Technology Education*, *18*(12), em2190. https://doi.org/10.29333/ejmste/12646

Alasgarova, G. A. (2022). A Rasch analysis of student coding attitudes survey. *The New Educational Review*, *67*, 237-248. https://doi.org/10.15804/tner.22.67.1.18

Angelo, T. A., & Cross, K. P. (2012). *Classroom assessment techniques*. Jossey Bass Wiley.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge.

Brookhart, S. M., & McMillan, J. H. (2020). *Classroom assessment and educational measurement* (p. 296). Taylor & Francis. https://doi.org/10.4324/9780429507533

Graue, M. E. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment*, *1*(4), 283-309. https://doi.org/10.1207/s15326977ea0104_1

Green, J., Liem, G.D., Martin, A.J., Colmar, S., Marsh. H.W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school : Key processes from a longitudinal perspective. *Journal of Adolescence, 35*(5), 1111-1122.https://doi.org/10.1016/j.adolescence.2012.02.016

Huang, L., Huang, F., Oon, P. T., & Mak, M. C. K. (2019). Constructs evaluation of student attitudes towards science. *Eurasia Journal of Mathematics, Science and Technology Education*, 15(12), em1792. https://doi.org/10.29333/ejmste/109168

Iseppi, L., Rizzo, M., Gori, E., Nassivera, F., Bassi, I., & Scuderi, A. (2021). Rasch model for assessing propensity to entomophagy. *Sustainability*, *13*(8), 1-21. https://doi.org/10.3390/su13084346

Kim, S., & Kyllonen, P. C. (2006). RASCH rating scale modeling of data from the standardized letter of recommendation. *ETS Research Report Series*, *2006*(2), 1-22. https://doi.org/10.1002/j.2333-8504.2006.tb02038.x

Mao, P., Cai, Z., He, J., Chen, X., & Fan, X. (2021). The relationship between attitude toward science and academic achievement in science: A three-level meta-analysis. *Frontiers in Psychology*, *12*, 1-12. https://doi.org/10.3389/fpsyg.2021.784068

Mertler, C. (2016). *Classroom assessment: A practical guide for educators*. Routledge. https://doi.org/10.4324/9781315266756

Mešić, V., Neumann, K., Aviani, I., Hasović, E., Boone, W. J., Erceg, N., ... & Repnik, R. (2019). Measuring students' conceptual understanding of wave optics: A Rasch modeling approach. *Physical Review Physics Education Research*, *15*(1), 010115. https://doi.org/10.1103/PhysRevPhysEducRes.15.010115

Muhammad, I., Angraini, L. M., Darmayanti, R., Sugianto, R., Usmiyatun, U., & da Silva Santiago, P. V. (2023). Students' ınterest in learning mathematics using augmented reality: Rasch model analysis. *Edutechnium Journal of Educational Technology*, 1(2), 89-99.

Oon, P. T., & Fan, X. (2017). Rasch analysis for psychometric improvement of science attitude rating scales. *International Journal of Science Education*, *39*(6), 683-700. https://doi.org/10.1080/09500693.2017.1299951

Oosterhof, A. (1999). *Developing and using classroom assessments*. Prentice-Hall Order Processing Center, PO Box 11071, Des Moines, IA 50336-1071.

Papanastasiou, E. C., & Schumacker, R. (2014). Rasch rating scale analysis of the attitudes toward research scale. *Journal of Applied Measurement*, 15(2), 189-199. https://doi.org/10.1037/t35505-000

Pell, T., & Jarvis, T. (2001). Developing attitudes to science scales for use with children of ages from five to eleven years. *International Journal of Science Education*, 23(8), 847-862. https://doi.org/10.1080/09500690010016111

Penuel, W. R., & Shepard, L. A. (2016). Assessment and teaching. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed., pp. 787–850). American Educational Research Association. https://doi.org/10.3102/978-0-935302-48-6_12

Pey Tee, O., & Subramaniam, R. (2018). Comparative study of middle school students' attitudes towards science: Rasch analysis of entire TIMSS 2011 attitudinal data for England, Singapore and the USA as well as as psychometric properties of attitudes scale. *International Journal of Science Education*, *40*(3), 268-290. https://doi.org/10.1080/09500693.2017.1413717

Rasch, G. (1960). *Probabilistic models for some intelligence and achievement tests*. Danish Institute for Educational Research.

Reckase, M. M. (1997). The past and the future of multidimensional item response theory. *Applied Psychological Measurement*, 21(1), 25-36. https://doi.org/10.1177/0146621697211002

Reid, C. A., Kolakowsky-Hayner, S. A., Lewis, A. N., & Armstrong, A. J. (2007). Modern psychometric methodology: Applications of item response theory. *Rehabilitation Counseling Bulletin*, 50(3), 177-188. https://doi.org/10.1177/00343552070500030501

Ryan, R. M., Connell, J. P., & Deci, E. L. (1985). A motivational analysis of self-determination and self-regulation in education. Research on motivation in education: *The Classroom Milieu*, *2*, 13-51.

Sabah, S., Hammouri, H., & Akour, M. (2013). Validation of a scale of attitudes toward Science across countries using Rasch model: Findings from TIMSS. *Journal of Baltic Science Education*, *12*(5), 692. https://doi.org/10.33225/jbse/13.12.692

Scoulas, J. M., Aksu Dunya, B., & De Groote, S. L. (2021). Validating students' library experience survey using rasch model. *Library and Information Science Research*, *43*(1), 1-12. https://doi.org/10.1016/j.lisr.2021.101071

Siegel, M. A., & Ranney, M. A. (2003). Developing the changes in attitude about the relevance of science (CARS) questionnaire and assessing two high school science classes. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *40*(8), 757-775. https://doi.org/10.1002/tea.10110

Soland, J. (2013). Predicting high school graduation and college enrollment: comparing early warning indicator data and teacher intuition. *Journal of Education for Students Placed at Risk (JESPAR)*, 18(3-4), 233-262. https://doi.org/10.1080/10824669.2013.833047

Steiner, R., & Sullivan, J. (1984). Variables correlating with student success in organic chemistry. *Journal of Chemical Education*, 61(12), 1072. https://doi.org/10.1021/ed061p1072

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement*, 22(4), 271-286. https://doi.org/10.1111/j.1745-3984.1985.tb01064.x

Sukri, A., Rizka, M. A., Purwanti, E., Ramdiah, S., & Lukitasari, M. (2022). Validating students' green character ınstrument using factor and Rasch model. *European Journal of Educational Research*, 11(2), 859-872. https://doi.org/10.12973/eu-jer.11.2.859

Suryadi, A., Fawaiz, S., Kurniati, E., & Swandi, A. (2023). Gender-and academic level-bias in MATS when measuring attitude towards science in Indonesia: A Rasch analysis. *Journal of Pedagogical Research*, *8*(1), 28-43. https://doi.org/10.33902/JPR.202422983

Veas, A., Castejón, J. L., Gilar, R., & Miñano, P. (2017). Validation of the Spanish adaptation of the *School Attitude Assessment Survey-Revised* using multidimensional Rasch analysis. *Anales de Psicología*, *33*(1), 74–81. https://doi.org/10.6018/analesps.33.1.235271

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

Zain, A. N. M., Samsudin, M. A., Rohandi, R., & Jusoh, A. (2010). Using the Rasch model to measure students' attitudes toward science in "low performing" secondary schools in Malaysia. *International Education Studies*, *3*(2), 56-63. https://doi.org/10.5539/ies.v3n2p56

Zeegers, P. (2004). Student learning in higher education: a path analysis of academic achievement in science. *Higher Education Research and Development*, *23*, 35-56. https://doi.org/10.1080/0729436032000168487