Academic Motivation Scale's psychometric attribute: Analysis using Rasch Measurement Model

Very Julianto[1*], Bambang Sumintono[2], Nandia Putri Zuhdi Almakhi[3], Hana Avetazain[4], Thifal Maida Wilhelmina[5], Destina Amelia Wati[6]

[1, 3, 4,5,6] Department of Psychology, Faculty of Social Sciences and Humanities, Universitas Islam Negeri Sunan Kalijaga, Indonesia
[2]Faculty of Education, Universitas Islam Internasional Indonesia, Indonesia

[*]Corresponding author: very.julianto@uin-suka.ac.id

Academic Motivation Scale's psychometric attribute: Analysis using Rasch Measurement Model

Abstract

In an academic context, it is important to measure students' academic motivation, which is closely related to learning improvement. One popular and widely used instrument to measure academic motivation is the 28-items Academic Motivation Scale (AMS) from Vallerand et al. (1989). This study aimed to evaluate the psychometric properties of the Indonesian version of AMS using the Rasch model measurement approach. A total of 1,193 Indonesian secondary school students participated in the study, and their self-reports were used to check instrument quality, including reliability, validity at the instrument level as well as at the item level. The findings showed that AMS Indonesia has adequate psychometric properties. However, four unfavorable items in one construct detected that they came from other dimensions, the rating scale used needs to be shortened, and a few items need to be revised because they have different item functioning. Further testing and refinement of the scale should be conducted.

Keywords: *Academic motivation; the Academic Motivation Scale (AMS); Vocational student; Rasch model; Indonesia*

Introduction

In the context of education, many previous studies have linked motivation to academic achievement and performance (Foong et al., 2021), learning creativity (Bhakti & Dwi Astuti, 2018), improved learning (Nordahl-Pedersen & Heggholmen, 2022), self-efficacy (Chen et al., 2013), finding meaning in life (Bailey & Phillips, 2015) and increased learning interest (González-Moreno, 2010). Students who are motivated by academics usually work hard to achieve their academic achievements (Mustafa et al., 2010; Steinmayr et al., 2019). Previous research shows that motivation will correlate with a person's academic achievement and accomplishments (Amrai et al., 2011; Bin Abdulrahman et al., 2023; Yousefy et al., 2012).

On the other hand, students with low academic motivation perform poorly in the learning process and often experience academic failure (Mauliya et al., 2020). Low academic motivation also makes students experience boredom in learning, so that their academic skills decrease. The more extreme the impact of low academic motivation is the increase in dropout rates, onset of mental health problems, and suicide due to poor academic performance (Fuertes et al., 2023; Orozco et al., 2018). These problems indicate that motivation has a positive relationship with a person's quality of life and has a major influence on the direction of life (Hopkins et al., 2021).

Regarding motivation significance, some researchers have developed instruments to measure this latent trait (McGill, 2012), for instance, in the education field, several instruments have been developed to measure motivation, such as the Problems in School Questionnaire (PIS) (Deci et al., 1981), Motivated Methods for Learning Questionnaire (MSLQ) (Pintrich et al., 1991), Student Motivation Scale (Martin, 2001), Questionnaire of Current Motivation (QCM) (Rheinberg et al., 2001), Flow Short Scale (FSS) (Rheinberg et al., 2003), Instructional Materials Motivation Survey (IMMS) (Keller, 2010), and Academic Motivation Scale (AMS) (Vallerand et al., 1989). The latter instrument, AMS, is regarded as a comprehensive tool that contains 28 items that assess students' intrinsic motivation, extrinsic motivation, and amotivation constructs, and has acceptable

reliability and validity (Marvianto & Widhiarso, 2019). According to Vallerand et al., (1992; 1993), theoretical background of AMS based on self determination theory developed by Deci and Ryan (1985) which explain role of intrinsic and extrinsic motivation in human behavior, where each individuals have basic psychological needs for autonomy, competence, and relatedness. AMS has been used at various levels of education, including in senior high school (Stover et al., 2012) and undergraduate (Javaeed et al., 2019). AMS has been used and adapted in various countries and languages. (Cody et al., 2021; Hopkins et al., 2021; Zhang et al., 2015; Cabras et al., 2023; Can, 2015; Kotera et al., 2023; Marvianto & Widhiarso, 2018).

Few studies have analyzed the psychometric features of AMS in Indonesia, and few studies have used the Rasch model measurement approach for vocational school students. The Rasch model provides comprehensive information on an instrument's psychometric attributes, including internal structure (construct validity), response process (targeting), and test content at the item level (Davis & Boone, 2021; Liu & Lim, 2020; Ratnaningsih et al., 2024). The current study aimed to analyze the psychometric properties of AMS among vocational students in Indonesia using the Rasch analysis approach, in particular the Rasch rating scale model (RSM).

Material and Methods

Instrumentation

The Indonesian version of the AMS was translated by Natalya (2018) into a short version. The full Indonesian version used in the present study was translated by Marvianto and Widhiarso (2019). There are seven constructs in the instrument: intrinsic motivation to know, intrinsic motivation to accomplish things, intrinsic motivation to experience stimulation, external regulation, introjected regulation, identified regulation, and amotivation*;* each construct consists of four items, with a total of 28 items. The response options for all items were the same, using a seven-point Likert-type scale ranging from strongly disagree (1) to neutral (4), and strongly agree (7). Except for the amotivation construct, the scores were reversed because all items were unfavorable statements. The instrument is self-administered and provided to respondents using an online platform (Google Forms), in which respondents can choose to participate at their preferred time.

Research Participants

This study was approved by the letter number B-2203.17/Un.02/L3/TU.00.9/06/2021 from the Chairman of Research and Community Services/LPPM. This study employed a quantitative approach using a cross-sectional survey design. The data were collected through an online questionnaire using a Google form platform across several provinces in Indonesia. The respondents for this study were vocational school students and young adults aged (16–18 years old), which is the age range of senior high school students in Indonesia (Adioetomo et al., 2014). They were invited through their schools' principals' permission, where they distributed the website link of the questionnaire using social media to students. Informed consent was obtained from the first page of the research scale. A prize of IDR 50,000 is up for grabs for 50 lucky participants through a lottery system, which includes an estimate of the time required to complete the scale. Participation was voluntary and was free of coercion. The school principal, the deputy head of curriculum and student affairs, and guidance and counseling instructors helped in data collection.

A total of 1381 respondents participated in the survey. This number is above the minimum requirement needed to represent vocational school students in Indonesia. The next step is extracting data from google form to

Microsoft excel, then preparing to code to do data cleaning and validation procedure using WINSTEPS version 3.73, a Rasch measurement model software, for detecting outliers (10 respondents who answered all minimum or maximum values) and misfit response (178 respondents with Outfit MNSQ index bigger than 2) (Andrich, D., & Marais, 2019; Widhiarso & Sumintono, 2016). As a result, 1193 respondents were analyzed further in this study for instrument empirical testing, and their demographic profiles are presented in Table 1.

[Insert Table 1]

Measurement Model

Achievement motivation, as a variable in this study, is usually referred to as the latent trait variable, which is not visible and measured using the AMS instrument. The item response theory (IRT) approach, in this case, the Rasch model, treats the measurement of latent variables as requiring the establishment and use of a linear equal-interval scale, as in physical sciences (Andrich & Marais, 2019; Boone et al., 2014). The IRT itself is grounded in psychometrics and focuses on the relationship between individuals' latent traits (i.e. atitude, perception or opinion) and their responses to test items using probabilistic model, which provide accuracy and precision measurement result of the trait (Sumintono & Widhiarso, 2014). In the case of the AMS instrument, which asks about respondents' perceptions or attitudes, it produces raw data that are ordinal in nature with a contaminated scaling problem. Then, the Rasch model needs to be used, where raw data are transformed with a probabilistic model and logarithm function that places both persons and items measures onto one common ruler to measure the construct (Bond & Fox, 2015; Liu & Lim, 2020).

This study used a Rasch rating scale model (RSM) approach because the data obtained from the AMS instrument are Likert-type with seven ratings, which are polytomous data (Andrich & Marais, 2019). Using WINSTEPS software, the raw data from the respondents were mathematically transformed into *logit* (logarithm odd unit), which resulted in three logit estimates: person location, item calibration, and an overall set of thresholds that inform whole instrument quality (Bond & Fox, 2015; Boone & Staver, 2020).

The results of Rasch analysis are comprehensive and involve information at the instrument level, up to item and person analysis, usually called individual-centered statistics (Engelhard & Wind, 2017). For instance, in terms of reliability, it can provide information on item and person reliability as well as item and person separation, not just relying on Cronbach's alpha, which can be problematic as it is a norm-referenced index that is sensitive to sample size (Boone et al., 2014). Rasch analysis emphasizes the unidimensional variable of the scale that makes factor analysis irrelevant; further, it is also informing, unexplained variance as an indication of the instrument's ability to separate a person's abilities based on how best the items were separated to cover different aspects of the latent variable being measured (Andrich, D., & Marais, 2019). At the item and person level, its validity is measured using several indices, which are infit and outfit mean squares (MnSqs) (acceptable range between 0.5 to 1.5), standardized residuals (ideal value is 0.0), and point measure correlations (PMC) (Andrich, D., & Marais, 2019; Boone, W. J., & Staver, 2020; Boone et al., 2014). The three indices indicate to assess how well the data (i.e., the responses of individuals to items) fit the expected mode, for instance item with MnSq less than 0.5 showing the response too predictable, whereas MsSq higher than 1.5 showing the item unpredictable (Bond & Fox, 2015).

Result

Reliability and Dimensionality

To determine whether the instrument had good reliability and validity at the instrument level, Table 2 shows the details of the findings. Reliability indices for person (0.72), item (1.00), and alpha (0.78) indicated that the consistency of person and item responses was 'very good' (Boone & Staver, 2020). Separation indices, which inform the grouping of respondents (1.62) and how widespread item difficulty level (21.75) showing the indices for both people (rounded to two) and items (more than three) is an acceptable value, supporting the fact that very reliable data and instruments were collected in this study.

[insert table 2]

The table above also shows that the data gathered fits the model as the outfit mean square value is close to one (an ideal value) for both person and item, which also confirms the significant value of the chi-square test result. In addition, unidimensionality which is the main requirement of the Rasch model, where all the items must measure a single construct, is assessed by looking at the fit statistics for the items and carrying out a principal components analysis (PCA) of the residuals (Andrich, D., & Marais, 2019; Bond & Fox, 2015). Two indices indicating this, which are raw variance, should be more than 20% and the eigenvalue of unexplained variance less than three (accounted for secondary dimension), one index as shown in Table 2, is in satisfactory level of unidimensionality of the AMS instrument (Fisher, 2007). For the eigenvalue index, it is more than three, indicating that around six items come from other dimensions when looking closely at all items coming from the same construct, which is amotivation.

Rating Scale Functioning.

The rating scale analysis used in the current study was a seven-type Likert rating agreement scale. In each rating category, the responses for rating 1 (strongly disagree) to 7 (strongly agree) were more than 100 times, and the value of Outfit MNSQ in each rating was less than 2, both of which were very acceptable and followed a minimum requirement (Linacre, 1999). The average measure score increases in the advancing categories, from -0.08 logit for the first rating (1) to +1.74 logit in the last rating (7), indicating that the seven ratings in this instrument followed the monotonic principle. This evidence shows that persons with higher agreement with the latent variable have higher probabilities of choosing higher ratings.

However, in terms of the ideal distance value between rating scales, which should be 1.40 to 5.0 logit (Liu & Lim, 2020), The distance of the Rasch-Andrich threshold between seven rating categories is not on an ideal value. For instance, there are two distances between the rating value that has a small value, which is between rating 4 and 5 (0.57) and between rating 5 and 6 (0.62). As shown in Figure 1 (on the left), not all ratings are separated to have their own peak, only four ratings have their peak (rating 1-2-6-7). This information indicates that the rating scale with seven categories is not well understood by the respondents when answering the items in this study (Schutte et al., 2021). The effectiveness of each unit rating scale is crucial for obtaining correct and optimal results (Linacre, 2012), and one of the alternatives to increasing good results is collapsing lower response categories. As shown in Figure 1 (on the right side), the improved picture emerged when the two

lower step category ratings collapsed, where each response category represented a clearly distinct portion of the underlying trait. However, the five new ratings also do not have their peak, indicating that the four ratings are better for this scale.

[insert figure 1]

Item Properties

The essential results of Rasch model analysis can provide information at the item level, usually called individual-centered statistics (Engelhard & Wind, 2017). Winsteps software provides a result that informs psychometric attributes at the item level. The logit value or estimation of item location (calibrations) shows that there are no outlier items in the AMS instrument; all the standard error values are very small, indicating that the item has very good precision (Boone et al., 2014; Adams et al., 2021).

Item fit statistic information usually involves three indices: mean squares (MnSqs), t-statistics (ZStds), and point measure correlation (PMCs); however, a larger volume of data (more than 500 respondents) makes the ZStds no longer relevant anymore to look (Engelhard & Wind, 2017; Lee et al., 2021). For infit and outfit MnSqs, the ideal value is 1, and the acceptable range is between 0.5 and 1.5 (Bond & Fox, 2015). It can be seen in the table that two items (items 21 and 26) are beyond the acceptable value for Infit Mnsqs, and four items (items 25 to 28) do not fulfill the category for Outfit Mnsq. In total, five items were categorized as misfit. Interestingly, four out of five misfit items come from the amotivation construct; at the same time, all of these items are also categorized as other dimensions in the principal components analysis. This indicates that an unfavorable statement from this construct tends to confuse respondents in this study.

Something interesting in Rasch analysis, which can provide accurate and precise measurement at the item level that has interval scale quality, the items can be grouped based on their logit value (item difficulty level). Using the information of item mean (which is fixed at 0.0 logit) and the value of item standard deviation, or SD, (0.83), this study divides items into four groups (Table 3). As shown in the table below, there are interesting findings regarding the item difficulty group.

[insert table 3]

Using mean and SD thresholds, there are four items (15%) in the very difficult-to-agree group (item logit $\geq$ mean+SD), all of which come from the same construction, which is amotivation. The explanation for this is that the students found that amotivation statements were not in their favor. In the second group, there were eight items (28%); interestingly, all items from the intrinsic motivation to experience stimulation (IMTE) construct belonged to this group, showing statements about this considered difficult to agree, as well as IMT3, INR3, and two items from identified regulation (ER).

For the third group, which was easy for the respondent to agree with, there were 12 items (42%) that came from three constructs, which is the majority of the items. The last group consisted of four items (15%) from three constructs in the very easy-to-agree statement-likely category (item logit $\leq$ mean-SD).

Person Response and Targeting

Other instrument qualities can be shown in terms of how the respondents react to the scale, which detects outlier and misfit responses (Liu & Lim, 2020). As explained in the methodology section above, 188 respondents dropped out of the study, accounting for 1.3% of the total respondents, a small number. The rest of the respondents, 1195 persons, are an acceptable fit that answers the AMS scale by self-reporting their perception.

The respondents' answers to the scale were later analyzed to see how well it matched ranges of person abilities with the item difficulty level (Boone et al., 2014). It is important to know if the scale has a problem in terms of the floor and ceiling effect, where respondents' answers tend to skew below (lower score) or above (higher score), which indicates that the scale is incorrect in terms of targeting.

[insert figure 2]

Evidence of targeting is shown in Figure 2, which is usually called a Wright map or an item-person map. The far-left column shows the range of person scores, informing from the lower score of latent variables at the bottom to the higher score on the top, whereas item difficulties in the three columns to the right (item symbolize with XX). The real item difficulty level of the AMS scale is in the middle (Measure), while on the left from the center (bottom p=50%) shows items with half probability exceeding the bottom category of the scale, while one column in the far right (top p=50%) shows items with half probability below the top category of the scale. Figure 2 shows a good right map where item difficulties are well targeted to personal abilities, where nearly all the ranges of person columns match with the three-item columns. Further checking the person logit, no person had minimum or maximum extreme scores, indicating a very good ceiling and floor effect of the scale in the study.

Differential Item Functioning (DIF)

Lastly, DIF analysis was conducted to examine item bias according to demographic variables collected in the study (gender, domicile, and type of school). However, only gender with a good proportion should be checked for DIF analysis. An item was considered to have DIF if it fulfills three indices, which are the DIF contrast value of less than −0.5, or more than 0.5, the Rasch-Welch t value of less than −2.0, or more than 2.0, and the Mantel-Haenszel p (Probability) value of less than 0.05 (Bond & Fox, 2015; Boone et al., 2014). However, there were items from the AMS instrument that potentially had DIF. For instance, Table 4 shows a DIF analysis of three indices based on gender; there were three items that Rasch-Welch t value outside the acceptable range (IMTE2, IMTE4, and IR4) and four items in the Mantel-Haenszel probability (IMTE2, IMTE4, IR4, and ER3), but only two items with DIF contrast values that are not acceptable (IMTE4 and ER3).

[insert table 4]

In Figure 2, which shows a DIF plot based on gender, which has two groups, male and female respondents, the pattern of response is nearly similar to all items, except very slightly different for the four items mentioned in Table 4. If the curved line tends to be above, it is difficult to agree with the respondent; for instance, item IMT2 (the first blue circle on the left) shows that males tend to disagree with the items compared

to female respondents. These findings indicate the unique response of the AMS instrument from the scale of the Indonesian study, where at least one item has DIF based on sex (IMTE4).

[insert figure 3]

Discussion

The objective of this study was to validate the Indonesian version of the AMS instrument with a sample of vocational secondary school students from Indonesia using the Rasch measurement model. AMS has been tested by many researchers in Europe (Ferreira Lopes et al., 2018; Tóth-Király et al., 2017), Asia (Natalya, 2018; Zhang et al., 2015), America (Cokley, 2015; Vallerand et al., 1992, 1993), and Oceania (Hopkins et al., 2021) with mixed result. Many studies have mostly employed the classical test theory approach, such as research on the adaptation of the Indonesian version of the academic motivation scale, which shows that this scale has good validity and reliability so that it can be used (Marvianto & Widhiarso, 2018). Thus, limited study has been conducted in the Indonesian context using Rasch analysis.

The findings of the present study indicate that the AMS instrument needs some revisions to become a valid tool for measuring achievement motivation in the Indonesian sample. Unidimensionality of the scale, which is based on principal component analysis of residuals, a primary dimension was above 60% of the variances in in-person responses showing good construct validity; however, with Eigen value 6.0, which shows that there is a second dimension in the scale and comes from one construct in which all items used unfavorable statements (Boone et al., 2014; Liu & Lim, 2020). This shows that negatively worded items in the AMS tend to have lower levels of discrimination and hand different types of responses than positively worded items because the items require cognitive processing complexity that causes participants to experience difficulty and affects their answers (Sliter & Zickar, 2014).

Research while all reliability indices are excellent, such as person and item reliability as well as the Cronbach's alpha index, indicating that the Indonesian sample in this study is large enough to have a good internal consistency of the scale. The index of item separation (21.75) clearly indicates that it has better item difficulty ordering even though the scale has 28 items, while with a person separation index round to two informed that the scale was able to differentiate two levels of personal abilities, a threshold criterion (Fisher, 2007). This result is similar to that of research that uses the classical test theory approach (Marvianto and Widhiarso, 2018).(Marvianto and Widhiarso, 2018).

Regarding the rating scale analysis, monotonic assumptions are fulfilled with the increase in the measure, and the Mn-Sq index also confirms this. However, in terms of step calibration, the findings from the Indonesian sample show that the 7-point scale employed needs to be collapsed to improve the rating functionality (Schutte et al., 2021). When two middle ratings are collapsed into five choices, the peak of each rating still needs further collapse (Yamashita, 2022). Therefore, in this case, four rating choices appear more appropriate for AMS to make the rating functional and clearer for the Indonesian samples. Rating scale analysis can be performed using the Rasch Model approach, which is able to provide definite answers regarding the validity of the instrument so that it can improve the quality of the instrument used (Liu & Lim, 2022; Schutte et al., 2021).Regarding the rating scale analysis, monotonic assumptions are fulfilled with the increase in the measure, and the Mn-Sq index also confirms this. However, in terms of step calibration, the findings from the Indonesian sample show that the

7-point scale employed needs to be collapsed to improve the rating functionality (Schutte et al., 2021). When two middle ratings are collapsed into five choices, the peak of each rating still needs further collapse (Yamashita, 2022). Therefore, in this case, four rating choices appear more appropriate for AMS to make the rating functional and clearer for the Indonesian samples. Rating scale analysis can be performed using the Rasch Model approach, which is able to provide definite answers regarding the validity of the instrument so that it can improve the quality of the instrument used (Liu & Lim, 2022; Schutte et al., 2021).

As the unique approach of the Rasch model emphasizes individual-centered statistics, a comprehensive investigation at the item level can be conducted. There were six items out of 28 that did not have good fit statistics indices, especially Outfit MnSqs, where four of them come from amotivation constructs; the rest of the AMS instrument item has good item psychometric attributes, all in the acceptable range for MnSqs, and no polarization since all point measure correlation indices were positive (Boone et al., 2014). The four items from the amotivation construct category as misfit showed that the item content confused some Indonesian samples, which was the item that had a higher logit and, at the same time, a bigger loading contrast in the unidimensionality test (Boone & Staver, 2020). Therefore, item refinement for this construct needs to be performed in future studies. Rasch analysis can detect and provide information at the individual level on the same scale (Adams et al., 2021).(Adams et al., 2021).

For person-item targeting, the findings show that it does not have a ceiling or floor effect, which is negligible from the total sample (Liu & Lim, 2020; Davis & Boone, 2021; For person-item targeting, the findings show that it does not have a ceiling or floor effect, which is negligible from the total sample (Liu & Lim, 2020; Davis & Boone, 2021; Stano et al., 2024)). The mean of the person measure (+1.02 logit) was higher than the item mean measure (0.00 logit), indicating that the AMS instrument for the Indonesian sample in this study tends to be more easily compared to personal abilities. The standard deviation of person logit (0.47) is slightly lower than that of item (0.83), suggesting better variation in item difficulty level.

Three demographic data were analyzed in this study, but only sex was analyzed for DIF analysis. The findings show that only two items have a bias (IMTE4 and ER3), and the three indices used to measure DIF are not within the acceptable range. For instance, IMTE 4 with the statement *"Because I feel very happy when reading about various interesting subjects"* is an item that is difficult for respondents to agree with. Students do not always focus on a particular interest; there are times when they feel saturated, even with things that interest them (Astalini et al., 2018). This means that these two items require revisions to make them fairer. DIF, which is an important part of the Rasch Model, is often used to obtain information related to psychometric measures because DIF can verify balance and fairness in the assessment process of participants with various demographic backgrounds and needs and detect item bias (Hagquist & Andrich, 2017; Humphry & Montuoro, 2021).

Educational psychologists can use the Academic Motivation Scale to measure students' academic motivation because it is valid and reliable. By measuring the condition of students' academic motivation, problems caused by low academic motivation, such as poor academic performance and achievement (Mauliya et al., 2020), mental health problems, school dropout rates, and suicide (Fuertes et al., 2023; Orozco et al., 2018), were identified. Moreover, motivation is an important aspect of an individual's life, as it is positively correlated with their quality of life (Hopkins et al., 2021).

The present study's limitation is that participants were recruited using an online survey and were vocational secondary school students, which means they may not be representative of the Indonesian students'

population. As a developing country, several regions of Indonesia are still not connected to the internet. The next limitation of this study is focused on internal psychometric attribute analysis using the Rasch measurement model, while other important findings, such as the external validity measure, could address the degree to which this motivation measure corresponds to any other theoretically related construct.

The Indonesian version of the AMS has adequate psychometric properties, especially in terms of construct validity, person and item reliability, item location, and item properties. Based on the Rasch model analysis in the present study, the following suggestions can be made to improve the test precision. First, a collapse rating, such as making it a four-point rating scale, is needed to make it functional. Second, all the items detected as misfit (for instance, in the amotivation construct) or having DIF should be revised. This scale refinement will make AMS has better functionality of AMS.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of Competing Interest

The research licensing process has been carried out so that it received approval from the ethics committee of the university-affiliated first author. The authors declare no conflict of interest.

References

Adams, D., Chuah, K.M., Sumintono, B. and Mohamed, A. (2021). Bricks to Clicks: Students' Engagement in E-Learning during the COVID-19 Pandemic. *Asia Pacific Journal of Educators and Education,* 36 (2), 99-117. http://dx.doi.org/10.21315/apjee2021.36.2.6

Adioetomo, S. M., Posselt, H., & Utomo, A. (2014). UNFPA Indonesia Monograph Series 2: Youth in Indonesia. In *UNFPA Indonesia Monograph Series* (Vol. 2, Issue July).

Amrai, K., Motlagh, S. E., Zalani, H. A., & Parhon, H. (2011). The relationship between academic motivation and academic achievement students. *Procedia - Social and Behavioral Sciences*, *15*, 399–402. https://doi.org/https://doi.org/10.1016/j.sbspro.2011.03.111

Andrich, D., & Marais, I. (2019). *A Course in Rasch Measurement Theory, Measuring in the Educational, Social and Health Sciences*. Springer.

Astalini, A., Kurniawan, D. A., & Putri, A. D. (2018). Identifikasi Sikap Implikasi Sosial dari IPA, Ketertarikan Menambah Waktu Belajar IPA, dan Ketertarikan Berkarir Dibidang IPA Siswa SMP Se-Kabupaten Muaro Jambi. *Tarbiyah : Jurnal Ilmiah Kependidikan*, *7*(2 SE-Articles). https://doi.org/10.18592/tarbiyah.v7i2.2142

Bailey, T., & Phillips, L. (2015). The influence of motivation and adaptation on students' subjective well-being, meaning in life and academic performance. *Higher Education Research & Development*, *35*, 1–16. https://doi.org/10.1080/07294360.2015.1087474

Bhakti, Y. B., & Dwi Astuti, I. A. (2018). The Influence Process of Science Skill and Motivation Learning with Creativity Learn. *Journal of Education and Learning (EduLearn)*, *12*(1), 30–35. https://doi.org/10.11591/edulearn.v12i1.6912

Bin Abdulrahman, K. A., Alshehri, A. S., Alkhalifah, K. M., Alasiri, A., Aldayel, M. S., Alahmari, F. S., Alothman, A. M., & Alfadhel, M. A. (2023). The Relationship Between Motivation and Academic Performance Among Medical Students in Riyadh. *Cureus*, *15*(10), e46815. https://doi.org/10.7759/cureus.46815

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

Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model: Fundamental measurement in the human sciences, 3rd ed. In *Applying the Rasch model: Fundamental measurement in the human sciences, 3rd ed.* Routledge/Taylor & Francis Group.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Rasch Analysis in the Human Sciences. *Dordrecht: Springer*. https://doi.org/http://dx.doi.org/10.1007/978-94-007-6857-4

Cabras, C., Konyukhova, T., Lukianova, N., Mondo, M., & Sechi, C. (2023). Gender and country differences in academic motivation, coping strategies, and academic burnout in a sample of Italian and Russian first-year university students. *Heliyon*, *9*(6), e16617. https://doi.org/https://doi.org/10.1016/j.heliyon.2023.e16617

Can, G. (2015). Turkish version of the Academic Motivation Scale. *Psychological Reports*, *116*(2), 388–408. https://doi.org/10.2466/14.08.PR0.116k24w5

Chen, M., Changhua, N., & Liao, J.-L. (2013). Correlations among Learning Motivation, Life Stress, Learning Satisfaction, and Self-Efficacy for Ph.D. Students. *The Journal of International Management Studies*, *8*(1), 157–162.

Cody, B., Kuyini, A. B., & Smith, L. (2021). Psychometric Analysis of the Academic Motivation Scale with Native Arab College Students. *Journal of College Student Development*, *62*, 591–606. https://doi.org/10.1353/csd.2021.0055

Cokley, K. (2015). A confirmatory factor analysis of the Academic Motivation Scale with Black college students. *Measurement and Evaluation in Counseling and Development*, *48*(2), 124–139. https://doi.org/10.1177/0748175614563316

Davis, D. R., & Boone, W. (2021). Using Rasch analysis to evaluate the psychometric functioning of the other-directed, lighthearted, intellectual, and whimsical (OLIW) adult playfulness scale. *International Journal of Educational Research Open*, *2*, 100054. https://doi.org/https://doi.org/10.1016/j.ijedro.2021.100054

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*(5), 642–650. https://doi.org/10.1037/0022-0663.73.5.642

Deci, E. L., & Ryan, R. M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. Berlin: Springer Science & Business Media. https://doi.org/10.1007/978-1-4899-2271-7

Dişlen Dağgöl, G. (2013). THE REASONS OF LACK OF MOTIVATION FROM THE STUDENTS' AND TEACHERS' VOICES. *The Journal of Academic Social Sciences*, *1*, 35–45. https://doi.org/10.16992/ASOS.13

Engelhard Jr., G., & Wind, S. (2017). *Invariant Measurement with Raters and Rating Scales: Rasch Models for Rater-Mediated Assessments* (1st ed). Routledge. https://doi.org/https://doi.org/10.4324/9781315766829

Farrokhi, F., & Esfandiari, R. (2011). A many-facet Rasch model to detect halo effect in three types of raters. *Theory and Practice in Language Studies*, *1*(11), 1531–1540. https://doi.org/10.4304/tpls.1.11.1531-1540

Ferreira Lopes, P., Silva, R., Oliveira, J., Ambrosio, I., Ferreira, D., Crespo, C., Feiteira, F., & Rosa, P. J. (2018). Rasch Analysis on the Academic Motivation Scale in Portuguese University Students. *NeuroQuantology*, *16*. https://doi.org/10.14704/nq.2018.16.3.1062

Fisher, W. P. (2007). Rating Scale Instrument Quality Criteria. *Rasch Measurement Transactions*, *21*(1), 1095.

Foong, C. C., Liew, P. Y., & Lye, A. J. (2021). Changes in Motivation and its Relationship with Academic Performance among First-Year Chemical Engineering Students. *Education for Chemical Engineers*, *38*. https://doi.org/10.1016/j.ece.2021.11.002

Fuertes, H. G., Evangelista Jr, I. A., Marcellones, I. J. Y., & Bacatan, J. R. (2023). Student Engagement , Academic Motivation , and Academic Performance of Intermediate Level Students. *International Journal of Novel Research in Education and Learning Vol.*, *10*(3), 133–149. https://doi.org/10.5281/zenodo.8037103

González-Moreno, P. A. (2010). Students' motivation to study music: The Mexican context. *Research Studies in Music Education*, *32*(2), 185–199. https://doi.org/10.1177/1321103X10384211

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

Hopkins, E. G., Lyndon, M. P., Henning, M. A., & Medvedev, O. N. (2021). Applying Rasch analysis to evaluate and enhance the Academic Motivation Scale. *Australian Journal of Psychology*, *73*(3), 348–356. https://doi.org/10.1080/00049530.2021.1904794

Humphry, S., & Montuoro, P. (2021). The Rasch Model Cannot Reveal Systematic Differential Item Functioning in Single Tests: Subset DIF Analysis as an Alternative Methodology. *Frontiers in Education*, *6*(November), 1–12. https://doi.org/10.3389/feduc.2021.742560

Javaeed, A., Asghar, A., Allawat, Z., Haider, Q., Mustafa, K. J., & Ghauri, S. K. (2019). Assessment of Academic Motivation Level of Undergraduate Medical Students of Azad Kashmir, Pakistan. *Cureus*, *11*(3), e4296. https://doi.org/10.7759/cureus.4296

Kotera, Y., Conway, E., & Green, P. (2023). Construction And factorial validation of a short version of the Academic Motivation Scale. *British Journal of Guidance & Counselling*, *51*(2), 274–283. https://doi.org/10.1080/03069885.2021.1903387

Lee W.L., Chinna, K. and Sumintono, B. (2021). Psychometrics assessment of HeartQoL questionnaire: A Rasch Analysis. *European Journal of Preventive Cardiology*, *28*(12), e1–e5.

https://doi.org/https://doi.org/10.1177/2047487320902322
Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement*, *3*(2), 103–122.
Linacre, J. M. (2012). *A user's guide to Winsteps Ministeps Rasch-model computer programs.*
Liu, V. Y. Y., & Lim, S. M. (2022). A psychometric evaluation of the brief resilience scale among tertiary students in Singapore. *Asia Pacific Journal of Education*, *42*(3), 464–477. https://doi.org/10.1080/02188791.2020.1845120
Martin, A. J. (2001). The Student Motivation Scale: A Tool for Measuring and Enhancing Motivation. *Journal of Psychologists and Counsellors in Schools*, *11*, 1–20. https://doi.org/DOI: 10.1017/S1037291100004301
Marvianto, R. D., & Widhiarso, W. (2018). Adaptasi Academic Motivation Scale (AMS) Versi Bahasa Indonesia. *Jurnal Psikologi UGM*, *4*(1), 87–95. https://doi.org/10.22146/gamajop.45785
Mauliya, I., Relianisa, R., & Rokhyati, U. (2020). Lack of Motivation Factors Creating Poor Academic Performance in the Context of Graduate English Department Students. *Linguists : Journal Of Linguistics and Language Teaching*, *6*, 73. https://doi.org/10.29300/ling.v6i2.3604
McGill, M. (2012). Learning to Program with Personal Robots: Influences on Student Motivation. *ACM Transactions on Computing Education (TOCE)*, *12*. https://doi.org/10.1145/2133797.2133801
Mustafa, S. M. S., Elias, H., Noah, S. M., & Roslan, S. (2010). A Proposed Model of Motivational Influences on Academic Achievement with Flow as the Mediator. *Procedia - Social and Behavioral Sciences*, *7*, 2–9. https://doi.org/https://doi.org/10.1016/j.sbspro.2010.10.001
Natalya, L. (2018). Validation of Academic Motivation Scale: Short Indonesian Language Version. *ANIMA Indonesian Psychological Journal*, *34*. https://doi.org/10.24123/aipj.v34i1.2025
Nordahl-Pedersen, H., & Heggholmen, K. (2022). What promotes motivation and learning in project management students? *Procedia Computer Science*, *196*, 791–799. https://doi.org/https://doi.org/10.1016/j.procs.2021.12.077
Orozco, R., Benjet, C., Borges, G., Moneta Arce, M. F., Fregoso Ito, D., Fleiz, C., & Villatoro, J. A. (2018). Association between attempted suicide and academic performance indicators among middle and high school students in Mexico: results from a national survey. *Child and Adolescent Psychiatry and Mental Health*, *12*(1), 9. https://doi.org/10.1186/s13034-018-0215-6
Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. The University of Michigan.
Ratnaningsih, I.Z., Prihatsanti, U., Prasetyo, A.R. and Sumintono, B. (2024), "Validation of the Indonesian version of the psychological capital questionnaire (PCQ) in higher education: a Rasch analysis", *Journal of Applied Research in Higher Education*, https://doi.org/10.1108/JARHE-10-2023-0480
Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein fragebogen zur erfassung aktueller motivation in lern- und leistungssituationen. [A questionnaire to assess current motivation in learning situations.]. *Diagnostica*, *47*(2), 57–66. https://doi.org/10.1026/0012-1924.47.2.57
Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). *Die Erfassung des Flow-Erlebens*. https://doi.org/10.23668/psycharchives.8590
Schutte, L., Negri, L., Delle Fave, A., & Wissing, M. P. (2021). Rasch analysis of the Satisfaction with Life Scale across countries: Findings from South Africa and Italy. *Current Psychology*, *40*(10), 4908–4917. https://doi.org/10.1007/s12144-019-00424-5
Sliter, K. A., & Zickar, M. J. (2014). An IRT examination of the psychometric functioning of negatively worded personality items. *Educational and Psychological Measurement*, *74*(2), 214–226. https://doi.org/10.1177/0013164413504584
Stano, F., Pellicciari, L., La Porta, F., Piscitelli, D., Angilecchia, D., Signorelli, M., Giovannico, G., Pournajaf, S., & Caselli, S. (2024). Rasch analysis of the forgotten joint score in patients with total hip arthroplasty. *Journal of Rehabilitation Medicine*, *56*(SE-Original Report), jrm15774. https://doi.org/10.2340/jrm.v56.15774
Steinmayr, R., Weidinger, A. F., Schwinger, M., & Spinath, B. (2019). The Importance of Students' Motivation for Their Academic Achievement – Replicating and Extending Previous Findings. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2019.01730
Stover, J., de la Iglesia, G., Boubeta, A., & Liporace, M. (2012). Academic Motivation Scale: adaptation and psychometric analyses for high school and college students. *Psychology Research and Behavior Management*, *5*, 71–83. https://doi.org/10.2147/PRBM.S33188
Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi)*. Trim Komunikata Publishing House.
Törmäkangas, K. (2011). Advantages of the Rasch measurement model in analysing educational tests: An applicator's reflection. *Educational Research and Evaluation*, *17*(5), 307–320. https://doi.org/10.1080/13803611.2011.630562
Tóth-Király, I., Gábor, O., Dombi, E., Jagodics, B., Farkas, D., & Amoura, C. (2017). Cross-cultural

comparative examination of the Academic Motivation Scale using exploratory structural equation modeling. *Personality and Individual Differences*, *106*, 130–135. https://doi.org/10.1016/j.paid.2016.10.048

Vallerand, R. J., Blais, M. R., Brière, N. M., & Pelletier, L. G. (1989). Construction et validation de l'échelle de motivation en éducation (EME). [Construction and validation of the Motivation toward Education Scale.]. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *21*(3), 323–349. https://doi.org/10.1037/h0079855

Vallerand, R., Pelletier, L., Blais, M., Brière, N., Senécal, C., & Vallieres, E. (1992). The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. *Educational and Psychological Measurement*, *52*, 1003. https://doi.org/10.1177/0013164492052004025

Vallerand, R., Pelletier, L., Blais, M., Brière, N., Senécal, C., & Vallieres, E. (1993). On the Assessment of Intrinsic, Extrinsic, and Amotivation in Education: Evidence on the Concurrent and Construct Validity of the Academic Motivation Scale. *Educational and Psychological Measurement - EDUC PSYCHOL MEAS*, *53*, 159–172. https://doi.org/10.1177/0013164493053001018

Van Zile-Tamsen, C. (2017). Using Rasch Analysis to Inform Rating Scale Development. *Research in Higher Education*, *58*(8), 922–933. https://doi.org/10.1007/s11162-017-9448-0

Widhiarso, W., & Sumintono, B. (2016). Examining response aberrance as a cause of outliers in statistical analysis. *Personality and Individual Differences*, *98*, 11–15. https://doi.org/10.1016/j.paid.2016.03.099

Yamashita, T. (2022). Analyzing Likert scale surveys with Rasch models. *Research Methods in Applied Linguistics*, *1*, 100022. https://doi.org/10.1016/j.rmal.2022.100022

Yousefy, A., Ghassemi, G., & Firouznia, S. (2012). Motivation and academic achievement in medical students. *Journal of Education and Health Promotion*, *1*, 4. https://doi.org/10.4103/2277-9531.94412

Zhang, B., Li, Y. M., Li, J., Li, Y., & Zhang, H. (2015). The Revision and Validation of the Academic Motivation Scale in China. *Journal of Psychoeducational Assessment*, *34*(1), 15–27. https://doi.org/10.1177/0734282915575909

Appendices

Tables

Table 1. Demographic Profile of Respondent (N=1193)

Demographics	Respondent	Percentage (%)
Gender		
Male	590	49.4
Female	603	50.6
Domicile		
Rural	1071	89.8
Urban	122	10.2
Type of Vocational School		
Public	1126	94.4
Private	67	5.6

Table 2. Summary Statistics of Person and Items

	Person	Item
N	1193	28
Outfit Mean Square		
Mean	1.06	1.06
SD	0.83	0.41
Separation	1.62	21.75
Reliability	0.72	1.00
Alpha Cronbach	0.78	
Chi-square (χ^2)	32179**	
Raw Variance	60.3%	
Unexplained variance Eigen value	6.0	

** $p < 0.01$

Table 3. AMS Item grouping based on item calibration (N=1193)

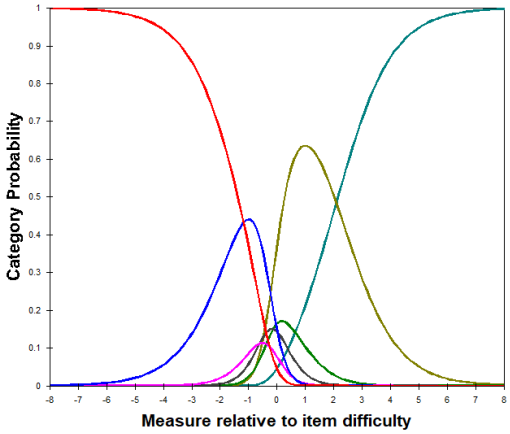
Construct	Difficulty level			
	Very easy to agree	Easy to agree	Difficult to agree	Very difficult to agree
intrinsic motivation to know (IMTK)	IMTK3	IMTK1, IMTK2, IMTK4		
intrinsic motivation to accomplish things (IMTA)		IMTA1, IMTA2 IMTA4	IMTA3	
intrinsic motivation to experience stimulation (IMTE)			IMTE1, IMTE2 IMTE3, IMTE4	
external regulation (IR)		IR1, IR2, IR3, IR4		
introjected regulation (INR)	INR1	INR2, INR4	INR3	
identified regulation (ER)	ER3, ER4		ER1, ER2	
Amotivation (AMO)				AMO1, AMO2, AMO3, AMO4

Table 4. Differential item functioning (DIF) of AMS based on gender (N=1193)

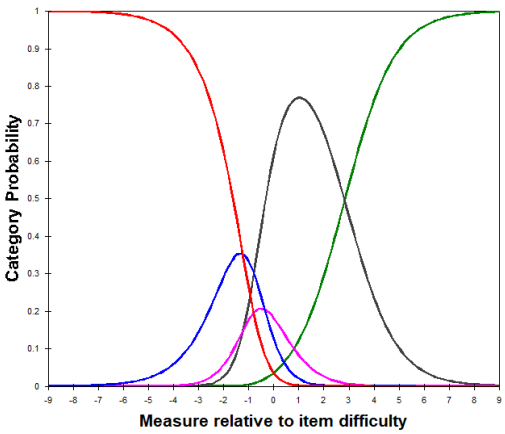
Item	**DIF Contrast**	**Rasch-Welch t**	**Prob.**
IMTE2	0.49	3.15	0.0048
IMTE4	0.59	3.71	0.0216
IR4	0.49	2.79	0.0035
ER3	-0.68	-1.05	0.0000

FIGURES

Figure 1. Category probability curves of the AMS instrument



AMS Original Rating Scale (1234567)



AMS with Collapsed Rating Scale (1233445)

Figure 2. Item-person map of the AMS instrument

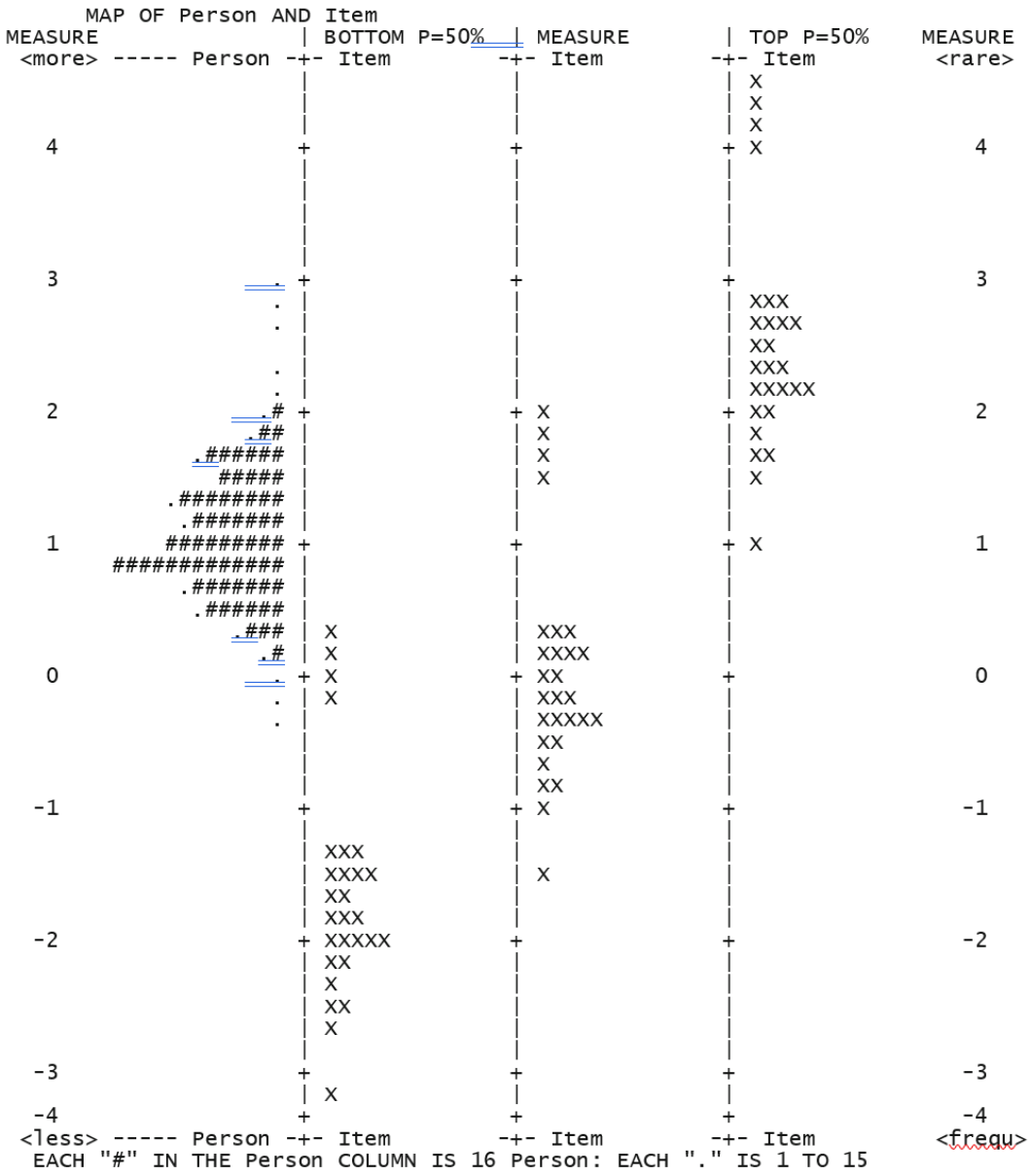
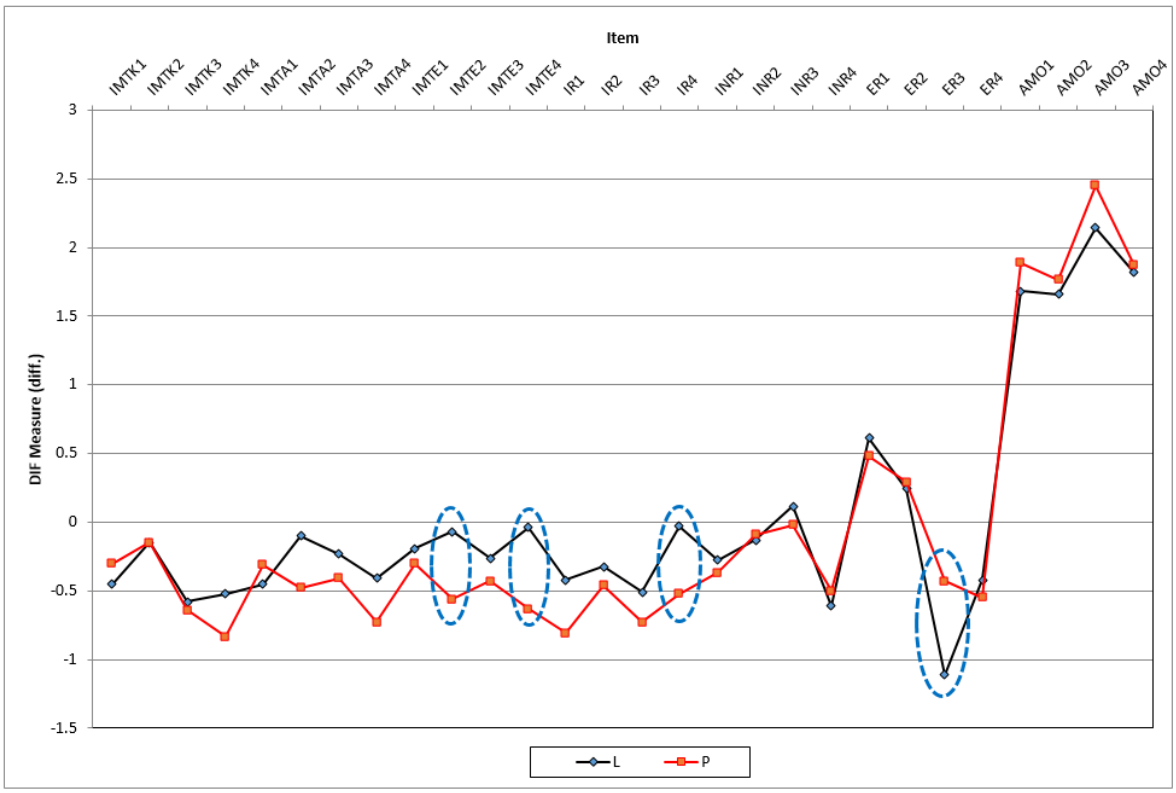


Figure 3. Different Item Functioning (DIF) of the AMS Instrument



Legend: L = male P = female