JPII 13 (3) (2024) 524-538

Jurnal Pendidikan IPA Indonesia

https://journal.unnes.ac.id/journals/jpii

ASSESSING PROSPECTIVE PHYSICS TEACHERS' INQUIRY SKILLS IN POST-PANDEMIC: RASCH ANALYSIS

Wahyudi[1,2], A. Setiawan[3*], A. Suhandi[4], A. Samsudin[5], Y. Kamin[6]

[1,3,4,5]Program of Science Education, Universitas Pendidikan Indonesia, Indonesia
[2]Program of Physics Education, IKIP PGRI Pontianak, Indonesia
[6]Schools of Education, Faculty of Social Science and Humanities, Universiti Teknologi Malaysia, Malaysia

DOI: 10.15294/jpii.v13i3.4015

Accepted: May 13th, 2024. Approved: August 31th, 2024. Published: August 31th 2024

ABSTRACT

Inquiry skills are necessary in the educational curriculum because they can help students construct scientific knowledge and skills, which are essential activities in science learning. This study aims to validate the scientific inquiry skills test using the Rasch Model parameters and to use Rasch analysis to evaluate the scientific inquiry skills of prospective physics teachers. This research used a quantitative method with a cross-sectional survey design. The adapted inquiry skills test questions consisted of 35 items covering nine aspects of inquiry skills. The participants in this research were 168 prospective physics teachers who were chosen randomly and came from educational institutions on the islands of Indonesia. The results show that the adapted inquiry skills test meets the validity and reliability criteria based on Rasch parameters. On average, students' skills in answering inquiry skills test questions are still poor. Analysis of differential item function reveals that the inquiry skills test items contain minimal biases based on university category and gender but are slightly larger based on semester level. The inquiry skills scores of students in semesters II, IV, and VI are not significantly different. The inquiry skills scores of male students are significantly different from those of female students. The adapted inquiry skills test instrument has met the validity, reliability, and slight bias criteria based on university category and gender. A small number of test items need to be corrected so they are not biased based on the student's semester level. The evaluation results show that male students show better inquiry skills than female students. Reorganization is needed in planning and implementing learning strategies that can facilitate the inquiry skills of prospective physics teachers.

Keywords: inquiry skills, prospective physics teachers, rasch analysis

INTRODUCTION

Education in science, technology, engineering, and mathematics (STEM) is a top priority for developing countries and a crucial aspect of 21st-century education (Gao et al., 2023; Surahman & Wang, 2023). Higher education provides an environment for people to support their future careers, including careers in science education (Wicaksono & Korom, 2023). In Indonesia, science lessons are the main subjects taught at the elementary and high school levels, specifically physics, chemistry, and biology. Various factors determine the quality of science education, but the most significant factor is science teachers' competence in teaching science (Fauth et al., 2019; Teig & Nilsen, 2022).

Science teacher preparation education must provide good content knowledge and practical scientific skills to produce high-quality science teachers (Morrell et al., 2020). Prospective science teachers must also have strong scientific inquiry skills (Özer & Sarıbaş, 2022) to become effective educators and motivate students for active learning. Inquiry skills for prospective science teachers, especially prospective physics teachers, not only help prospective physics teachers to understand physics concepts more profoundly but also develop critical and analytical thinking, design effective learning practices, solve problems,

*Correspondence Address
E-mail: agus_setiawan@upi.edu

communicate scientifically, and develop leadership and collaboration skills (Lu et al., 2021; Novitra et al., 2021).

The scientific inquiry skills used in contexts of the early years can vary. However, it refers to similar skills: observing, predicting, checking, recording, and communicating (Connor & Rosicka, 2020). According to Wenning (2007), not all inquiry processes are experimental, and sometimes, evidence and logic alone will be used to draw scientific conclusions. Scientific investigations based on observations will differ significantly from scientific investigations based on experiments. Aspects of scientific inquiry skills that are relevant and accommodate various methods of scientific inquiry include 1) identifying the problem and formulating a hypothesis, 2) making predictions, 3) designing experimental procedures, 4) conducting scientific experiments, 5) performing analysis and interpretation of data, 6) applying numerical methods and statistics, 7) explaining the unexpected results, and 8) utilizing available technology (Saputra et al., 2019). These aspects of scientific inquiry skills are relevant and suitable for measuring the scientific inquiry skills of prospective physics teachers.

Inquiry skills can be trained through inquiry-based learning. The results of the meta-analysis conducted by Firman et al. (2019) of 15 studies from various countries show that inquiry-based learning can be used in developing science learning and improving students' inquiry skills at every level of education. The inquiry learning model is based on search and discovery through a systematic thinking process. Inquiry-based learning makes students active, independent, and responsible for their learning (Spronken-Smith & Walker, 2010). Referring to levels of inquiry (Wenning, 2010), inquiry teaching practice is a series of inquiry spectrum learning starting from the stages of rudimentary skills until advanced skills.

Through the inquiry learning process, students have a better opportunity to engage with phenomena, understand the meaning of doing science, develop epistemological awareness of the nature of science, and develop inquiry skills (Constantinou et al., 2018). Inquiry skills will help build relevant knowledge in various sectors of society, social responsibility, and solving real problems in everyday life (Purwati et al., 2021). Inquiry skills have become an essential standard in the educational curriculum for prospective science teachers because they can help students construct scientific knowledge and laboratory skills, which are crucial activities in science learning (Setiono et al., 2019).

Even though much research reveals that inquiry-based learning can improve scientific inquiry skills, several obstacles are found in the implementation, mainly from the students' self-environment (Wenning & Vieyra, 2020). Moreover, there are many learning obstacles during the Covid-19 pandemic. All teaching processes were done via distance learning (Bahasoan et al., 2020), including at university levels worldwide (Altbach & De Wit, 2020). Even though this method is considered effective in suppressing the spread of the COVID-19 virus, distance learning requires much support in the form of technical support, emotional support, distance learning practices, online resources, and practical ways to move forward (Fackler & Sexton, 2020). However, not all educational institutions can provide maximum learning support, especially when replacing practical learning activities.

Online learning practices report many problems, such as limited teaching facilities and media, differences in facilities and teacher competence, limited internet access, and students not being serious about learning (Ezra et al., 2021). Literature has shown that poor audio and video quality is one of the daunting challenges for lecturers and prospective science teachers in implementing distance science learning (Murphy, 2009). Previously, student science teachers could complete hands-on activities in the classroom or laboratory to help build understanding through active learning experiences. However, after quarantine is implemented and teaching shifts to online, various challenges are found because learning emphasizes inquiry activities (Graham & Tolar, 2020). For prospective physics teachers, online learning cannot entirely replace inquiry activities in the laboratory. Lecturers and students must also develop more agility in adapting to online learning strategies (Koh & Daniel, 2022). The various problems and limited resources that arise increase obstacles in practicing scientific inquiry skills, especially for prospective science and physics teachers. So teachers must be able to choose the appropriate online learning mode (synchronous, asynchronous, and mixed) (Syawaludin et al., 2022) and appropriate learning media support because it will affect students' inquiry skills.

After the current Covid-19 pandemic, learning activities for prospective teachers have returned to normal. Learning success needs to be evaluated, especially prospective physics teachers' scientific inquiry skills competency. Assessing the scientific inquiry skills of prospective physics teachers is an essential step in preparing them to become competent and effective educators. By ensuring that scientific inquiry skills are emphasized in the education of student physics teachers, we can ensure that the next generation of physics educators will be more skilled and competent in teaching physics. This evaluation also allows prospective teachers to plan their personal development in scientific inquiry skills, which can improve their ability to teach physics with a more active and inquiry-based approach.

No research has used objective measurements such as Rasch analysis to evaluate the scientific inquiry skills of prospective physics teachers in Indonesia, especially when the COVID-19 pandemic ends. The impact of the COVID-19 pandemic on students' scientific inquiry skills can also be an evaluation of the effectiveness of physics learning during the pandemic. It is important to remember that the impact of the COVID-19 pandemic may differ in different regions and educational institutions, and differences in the scientific inquiry skills of prospective physics teachers during and after COVID-19 may vary. A comprehensive evaluation and measurement needs to be done to understand more deeply the differences in scientific inquiry skills that may occur.

Previous research has developed and measured students' inquiry skills, such as the scientific Inquiry literacy Test (ScInqLiT) (Saputra et al., 2019; Wenning, 2007), closed question test to measure students' scientific inquiry skills (Čipková & Karolčík, 2018), independent skills observation sheet (Elisanti, 2020) and taxonomy of inquiry skills for experimental activities (Ješková et al., 2018). However, some of these studies still use classical measurement theory in developing and measuring students' scientific inquiry skills. Using and developing measurement instruments is an integral part of the history of science education research (Liu, 2020), so objective analytical methods are needed in producing instruments for measuring scientific inquiry skills. The Rasch model provides a useful methodological tool for investigating the validity and reliability of measurements, allowing researchers to collect data from several observed indicators to express the results as a single variable (Lamprianou, 2020).

Research that uses Rasch Analysis as an objective measurement to develop instruments and measure inquiry skills is carried out by Kuo et al. (2015) by creating a Multimedia-based Assessment of Scientific Inquiry Abilities (MASIA), which includes skills of questioning, experimenting, analyzing, and explaining. Then, Lou et al. (2015) have developed the Inquiry Skills Assessment for Earth Science (iSA–Earth Science) to measure students' scientific inquiry skills in earth science content, including the skills of (1) identifying questions for scientific investigation, (2) designing scientific investigations, (3) using tools and techniques to collect data, (4) analyzing and describing data, (5) explaining results and drawing conclusions, and (6) recognizing alternative explanations and predictions. Arnold et al. (2018) have also developed an instrument of Scientific Inquiry Competence (SIC) for scientific inquiry competency, including skills in formulating questions/hypotheses, designing experiments, and analyzing data. All scientific inquiry skill instruments have been developed to measure the scientific inquiry skills of secondary school students.

Studies evaluating the inquiry skills of prospective physics teachers in Indonesia are still limited. To the best of the author's knowledge, no research has been conducted to measure and evaluate scientific inquiry skills for prospective physics teachers in terms of gender, semester level, and university category using Rasch model analysis. Research shows that male and female students often have different interests and levels of achievement in science (Hoffmann, 2002; Sagala et al., 2019). However, science learning (including physics material) should not be biased differently towards male or female students, as recommended by the 2014 NRC (National Research Council, 2014). Differences in knowledge and skills do not only occur in physics lessons but also in mathematics lessons (Egara, 2023). Apart from that, research shows no differences in the scientific inquiry skills of school students based on level (Ješková et al., 2018). However, in the education of prospective science/physics teachers, the curriculum settings mean that students' inquiry skills can differ at the initial and upper semester levels. This research also measures the scientific inquiry skills of prospective physics teachers based on university categories at the end of the COVID-19 pandemic.

By applying Rasch analysis in the context of scientific inquiry skills, researchers can investigate the scientific inquiry skills of prospective physics teachers with precision measurements, map items on the inquiry skills test based on the level of difficulty, and investigate differential item functions based on various respondent demographic variables such as gender, semester level, and university categories. The scientific inquiry skills test produces dichotomous data, which can be assessed according to the difficulty level of each item and the student's skills to make it easier to evaluate instruments and students' abilities to answer the scientific inquiry skills test. The research results can describe the scientific inquiry skills of prospective physics teachers after the COVID-19 pandemic ends. It will illustrate the impact of online learning conditions during the pandemic on students' inquiry skills. The results of this research can also be a consideration for continuing online learning even though the pandemic has ended.

This research aims to assess the Indonesian adapted version of the scientific inquiry skills test by determining its validity and reliability to evaluate the scientific inquiry skills of prospective physics teachers in Indonesia and to classify the level of students' scientific inquiry skills based on gender and semester level. The main problem in this research is the characteristics of the inquiry skills instrument and the results of the inquiry skills evaluation of prospective physics teachers using Rasch Modeling analysis. This research is the first to provide an in-depth understanding of the quality of the adapted instruments and evaluate prospective physics teachers' inquiry skills based on differences in university category, gender, and semester level through objective measurements based on Rasch Modeling.

METHODS

This research used a cross-sectional survey design with quantitative methods (Creswell & Guetterman, 2019). A random sampling technique was used to collect data from 168 students in the physics education program from universities on the large islands of Indonesia. Students' anonymity was guaranteed to protect their identity. The participants were given 45 minutes to complete the online inquiry skills test under lecturer supervision in face-to-face or online classes. The demographic profile of participants is presented in Table 1.

Table 1. Demographic profile of participants (N=168)

Demography	Code	N	Percentage (%)
University Category			
State University	S	94	55.95
Private University	P	74	44.05
Gender			
Male	M	28	16.67
Female	F	140	83.33
Level of Semester			
Semester II	2	84	50.00
Semester IV	4	48	28.57
Semester VI	6	36	21.43

This study adapted the scientific Inquiry Skills Test (ScInqLiT) version 1.2 proposed by Wenning (2007). Wenning's ScInqLiT instrument is an instrument to measure high school students' inquiry skills and is in the form of a multiple choice with 35 questions. However, in this research, the test questions were considered relevant for prospective physics teachers because they had similar levels of logical thinking about abstract ideas. The inquiry skills aspect refers to the inquiry stages as in Table 2.

Table 2. Aspects of scientific inquiry and item distribution in the Scientific Inquiry Skills Test

Part of Instrument	Description	Item Code
Students' Demographic	University Category: State (S), Private (P)	S/P
	Gender: Male (M), Female (F)	M/F
	Level of Semester: Semester II (2), Semester IV (4), Semester VI (6)	2/4/6
Aspects of scientific inquiry	Identify a problem to be investigated;	Q13, Q17, Q22, Q35
	Use induction, formulate a hypothesis or model incorporating logic and evidence;	Q11, Q12, Q21, Q28
	Use deduction and generate a prediction from the hypothesis or model;	Q20, Q33, Q34
	Design experimental procedures to test the prediction;	Q1, Q4 Q9, Q24
	Conduct a scientific experiment, observation, or simulation to test the hypothesis or model;	Q2, Q3, Q5, Q31
	Collect meaningful data, organize, and analyze data accurately and precisely;	Q10, Q14, Q18, Q19, Q23
	Apply numerical and statistical methods to numerical data to reach and support conclusions;	Q25, Q26, Q29, Q30
	Explain any unexpected results;	Q7, Q8, Q15, Q16
	Use available technology, report, display, and defend the results of an investigation.	Q6, Q27, Q32

The adaptation of inquiry skills instruments included language translation into Indonesian, adjusting images/graphs, and replacing relevant questions with inquiry skills indicators. Supervisors and senior physics education lecturers validated the instrument adaptation. The correct answer is given one point, and the wrong answer is given no points. Thus, respondents who answered all answers correctly received a maximum score of 35 points.

Data was collected from universities in several large provinces in Indonesia. Permission was also requested from the study program and course lecturers to distribute the test to prospective physics teachers. Data collection was carried out online using Google Forms. Students could access the test inquiry skills online using standard browser applications such as Mozilla Firefox, Google Chrome, Microsoft Edge, and others. Students were allowed to answer questions for 45 minutes, including filling in demographic data before filling in the inquiry skills test questions. Test answer data was recorded in Google Forms and saved in .xls (Ms. Excel). The researchers collaborated with lecturers to give, observe, and monitor students' completion of inquiry skills tests.

Descriptive and comparative data analysis used open-source JASP software version 0.17.2.1 (JASP Team, 2023). Meanwhile, Rasch Model analysis used Winstep software version 5.4.3.0 (Linacre, 2023). Rasch analysis included performing Rasch modeling using joint Maximum Likelihood Estimate (JMLE), converting person values into a logit scale (interval data), starting from negative infinity to positive infinity (Boone et al., 2014). Next, this research presents descriptive statistics, reliability and validity analyses, Item-fit Analysis, Person-fit Analysis, Scalogram of Responses, Person Diagnostic of Misfit, DIF (differential item functioning) Analysis, Wright (Item-Person) Map, Item difficulty, and student ability classification.

The validity of person and item matches is identified according to the infit-outfit mean square (MNSQ) mean, where the acceptable range is from 0.5 to 1.5. However, 1.6 is still considered acceptable (Soeharto & Csapó, 2022), and point measure correlation (PTM CORR) is positive (Boone et al., 2014). The 1.00 logit value of the MNSQ infit-outfit is the ideal value of the suitability criteria in the validity of Rasch analysis (Andrich, 2018). The accepted ZSTD infit-outfit as a validity parameter is -2.0 to +2.0 for samples of less than 250 (Azizan et al., 2020; Boone & Staver, 2020). However, the ZSTD and PTM CORR parameters can be ignored if the MNSQ parameters are acceptable (Linacre, 2023; Soeharto & Csapó, 2021). The construct validity of the inquiry skills test is confirmed through unidimensionality analysis in Rasch analysis. Construct validity in question refers to the criteria of whether the items (inquiry skills test) are effective in measuring a range of respondents' abilities and whether the items can measure one variable comprehensively (Laliyo et al., 2022). Data can fundamentally be one-dimensional if the percentage of raw variance is at least 20% (for dichotomous data) and the percentage of unexplained variance first contrast is a maximum of 15% (Linacre, 2023; Sunjaya et al., 2022).

The DIF analysis in this research compares all ability levels from two or more groups (university category, gender, and semester level). Significant DIF can be assessed based on two categories: probability significant ($p < 0.05$) using Welch's t-test (Khalaf et al., 2022) and DIF contrast of more than 1.00 logits (Fan et al., 2022). Wright's map is presented to confirm the targeting criteria between item and person. DIF analysis (differential item functioning) evaluates item bias using the testing method. All Rasch testing procedures refer to Bond et al. (2021) and Linacre (2023). Analysis of the Logit Value of Person (LVP) is used to classify students' abilities to answer inquiry skills test questions. There are limits to classifying students' level of ability in answering inquiry skills test questions: LVP $\geq$ M+SD, M $\leq$ LVP < M+SD, M-SD $\leq$ LVP < M, and LVP < M-SD (Adams et al., 2021).

Data normality and homogeneity tests were analyzed as a prerequisite for comparative tests of scientific inquiry skills based on university category, gender, and semester level. Because the criteria for normality and homogeneity of the data were not met, a comparative test of scientific inquiry skills based on university categories used the Whitney Test. Meanwhile, the comparative test of scientific inquiry skills based on semester-level categories used the Kruskal-Wallis Test (Goss-Sampson, 2020; King et al., 2018).

RESULTS AND DISCUSSION

In this research, quantitative data is analyzed to validate the inquiry skills test using the Rasch model by finding out whether the data follows the Rasch model measurements, to evaluate the inquiry skills of prospective physics teachers, and to find out whether there are test questions that function differently for the university category, gender, and semester level. The results of data analysis for this purpose are reported in the following section. Validity and reliability are the main parameters in Rasch's analysis. A summary of statistical data from the results of Rasch's analysis on student inquiry skills is presented in Table 3.

Table 3. Summary statistics of person and items

Psychometrics attribute	Person	Item
Number (N)	168	35
Measure (logit)		
Mean	-0.71	0.00
SD, Standard Deviation	0.53	0.66
SE, Standard Error	0.04	0.11
Mean INFIT		
MNSQ	1.00	1.00
ZSTD	-0.04	0.01
Mean OUTFIT		
MNSQ	1.02	1.02
ZSTD	0.00	-0.06
Separation	0.91	3.52
Reliability	0.45	0.93
Cronbach's Alpha	0.48	
Uni-dimensionality		
Raw variance explained by measures	12.7%	
Unexplained variance first contrast	8.9%	

JMLE estimates in Rasch analysis of dichotomous data are used to validate the student inquiry skills test (35 items, 168 persons), adapted for Indonesia. Item and person parameters are used to validate the inquiry skills test of prospective physics students. Based on the predetermined criteria and the results of data analysis (Table 3), both the person and the items in the inquiry skills test have met the valid criteria. The quality analysis results of the inquiry skills test items show that the MNSQ ranges from 0.85 to 1.28 with a standard error of less than 0.5, so the instrument meets the fit and valid criteria.

The percentage of observations from raw variance and unexplained variance first contrast is presented in Table 3. The results of the analysis show that the percentage of raw variance of the inquiry skills test is only 12.7%. Even though it has not reached the acceptable threshold (20%), the inquiry skills test items are quite effective in measuring students' inquiry skills. Eigenvalue percentages from unexplained variance reinforce this first contrast, amounted to 8.9%, and met the minimum criteria of 20%. Thus, the inquiry skills test items fulfill construct validity, which effectively measures respondents' inquiry skills and can measure one variable comprehensively. The results displayed in Table 3 show that the item separation index is classified as high (≥ 3 logits), and item reliability is also classified as high (≥0.67). It shows that prospective physics teachers are adequate in confirming the difficulty level of the inquiry skills test and strengthening the construct validation of the inquiry skills instrument. Separation index on items >3 logit also shows that the item difficulty index can be representatively divided into high, medium, and low categories.

Results of investigations on person outfit MNSQ show a Rasch parameter ranges from 2.14 to 0.74, and the standard error per person is less than 0.5. Within this range, out of 168 people, two people do not meet the MNSQ criteria (in the MNSQ range of 0.5 to 1.6). The separation index of a person amounting to 0.91 (Table 3) shows that the test items can only differentiate between high and low student skills. Person reliability only produces an index in the low category, which means that the respondent's skills are not diverse enough. The person reliability index is proportional to the index of Cronbach's Alpha, which is estimated to describe the reliability of interactions between 168 respondents and 35 items of inquiry skills tests completely.

Table 4. Person-fit Analysis: Misfit Order

Entry No.	Person	Total Score	Total Count	JMLE Measure	Model S.E.	Outfit		PTM CORR
						MNSQ	ZSTD	
16	SF2	5	35	-1.94	0.49	**2.14**	**2.10**	-0.23
52	SF2	12	35	-0.74	0.37	**1.70**	**3.21**	-0.34
160	PF6	7	35	-1.51	0.44	**1.60**	1.62	-0.23
59	SF2	7	35	-1.51	0.44	**1.59**	1.60	-0.23
60	SF2	8	35	-1.33	0.42	**1.58**	1.77	0.00
25	SF2	14	35	-0.45	0.36	1.33	**2.07**	-0.04
107	PM6	16	35	-0.45	0.36	1.33	**2.06**	-0.01
108	PM6	16	35	-0.20	0.36	0.74	**-2.29**	0.63

Global Fit/Chi-Square: 5969,02; d.f: 5846; p=0.0000

Table 4 displays 8 data Person-fit Analyses, which are selected from 168 respondent (person) data and is data that does not meet the three acceptance criteria of MNSQ, ZSTD, and PTM CORR. However, the ZSTD and PTM CORR parameters can be ignored if the MNSQ parameters are acceptable. The data in Table 4 is sorted based on the highest to lowest MNSQ logit outfit. From the 8 data, only two people (1.19% of 168 people), 16-SF2 and 52-SF2, do not meet the MNSQ, ZSTD, and PTM CORR criteria; the rest still meet one or two criteria.

A misfit person can also be identified through person diagnostic in Rasch analysis. Figure 1 shows the diagnostic results of respondents of 16SF2 and 52SF2. There was not a single easy category item answered correctly by both respondents. However, several items were difficult for both respondents to answer correctly (unexpected success). Then, the evaluation of student inquiry skills was also done by analyzing students' inquiry skills achievements based on gender and semester level, as presented in Table 5.

Analysis of Logit value of person (LVP) classifies students' level of ability in answering inquiry skills test questions, including LVP ≥ M+SD, M £ LVP < M+SD, M-SD £ LVP < M, and LVP < M-SD. The results of the student inquiry skills data show that the mean logit of a person is -0.71 and the SD logit of person is 0.53, so students' inquiry skills are grouped into four categories: very high, high, low, and very low (Table 6). In detail, in the logit measure, the classification of students' skills in the inquiry skills test is also displayed on the Wright Map (Figure 2).

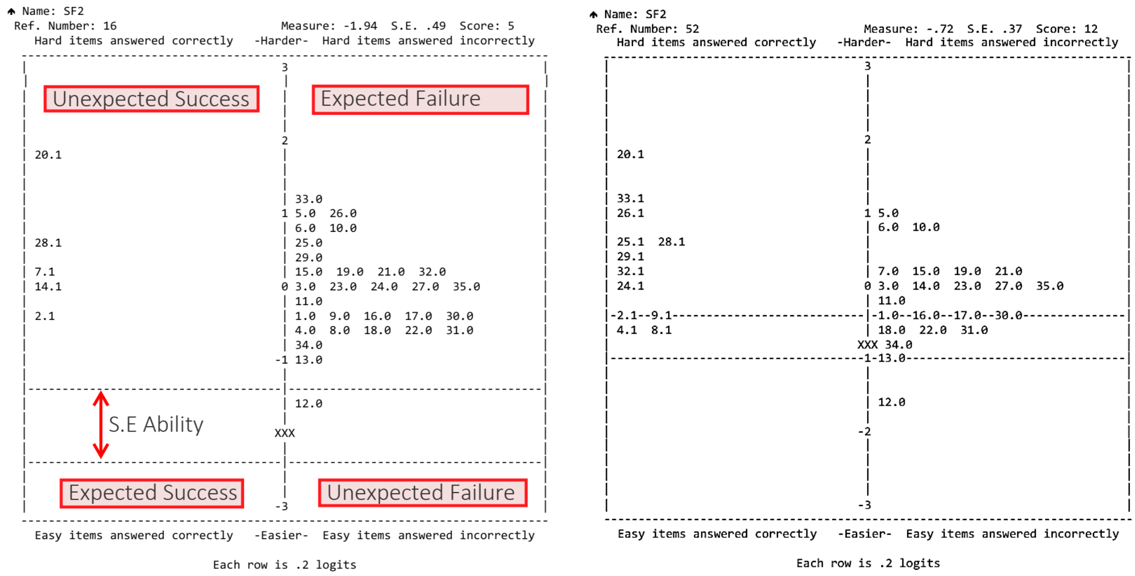

Figure 1. Person Diagnostic of misfit person (16SF2 & 52SF2)

Table 5. Descriptive Statistics of Students' Inquiry Skills

Data		N	Mean*	Min	Max	STDV
Gender	Male	28	14.32	8	30	5.29
	Female	140	11.81	4	22	3.19
Level of Semester	Semester II	84	12.54	4	26	3.78
	Semester IV	48	11.58	5	30	4.10
	Semester VI	36	12.36	8	22	3.01

*Maximum score = 35

Analysis of student abilities by gender (Table 6) finds that 19 of 28 (67.86%) male students and 75 of 140 (53.57%) female students are at a very high and high level of ability in the inquiry skills test. At low and very low ability levels, 9 out of 28 (32.14%) male students and 65 out of 140 (46.43%) female students. These results indicate that the proportion of male students with a high level of inquiry skills dominates in this research sample. Then, almost half of female students are at an average level of inquiry skills. Based on semester level, it is found that 48 of 84 (57.14%) second-semester students, 29 of 48 (60.42%) fourth-semester students, and 17 of 36 (47.22%) sixth-semester students are at a very high and high level of inquiry skills. The remaining 42.86%, 39.58%, and 52.78% are at low and very low levels at the semester level, respectively.

These results show that more than half of the second and fourth-semester students have very high and high levels of inquiry skills, while most sixth-semester students have low and very low levels.

Table 6. Categorization of the level of students' inquiry skills

Demographic	Students' ability level in inquiry skills test				Total
	Very high	High	Low	Very Low	
	LVP ≥ -0.18	-0.71 ≤ LVP < -0.18	-1.24 ≤ LVP < -0.71	LVP < -1.24	
Gender: Male	8	11	7	2	28
Female	17	58	44	21	140
Level of Semester: II	17	31	26	10	84
IV	3	26	15	4	48
VI	5	12	10	9	36
Total N (%)	25 (14.88%)	69 (41.07%)	51 (30.36%)	23 (13.69%)	168

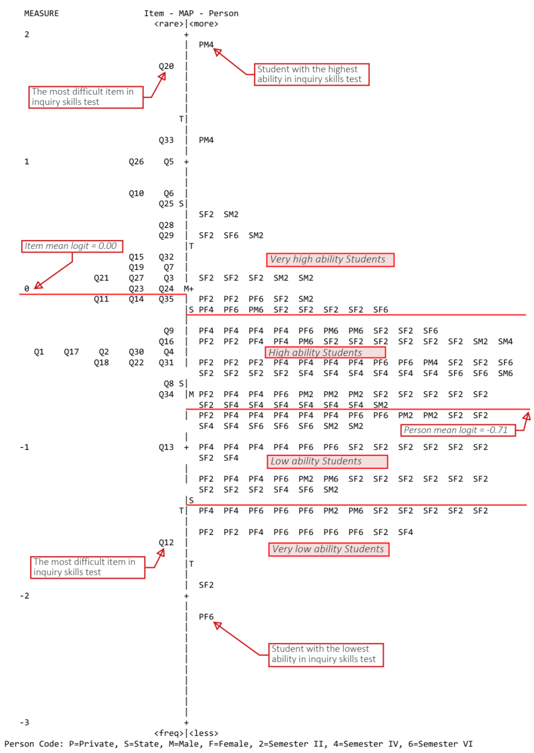

Figure 2. Wright map of person in inquiry skills test

The Rasch model has scaled the inquiry skills of prospective physics teachers from negative infinity to positive infinity, with 0 logits being the average measure of student inquiry skills. The Wright map shows that the person logit average (-0.71) is lower than the item logit average (0.00). It shows that, generally, students still have a low ability to answer test questions. Students with the code PM4 (specifically, entry number 104, private university, male, semester IV) have the highest level of ability in the inquiry skills test. The student can answer the inquiry skills test questions correctly, 26 out of the 35 questions given (74.29%). Students with the code PM6 (specifically, entry number 163, private university, male, semester VI) have the lowest ability level in the inquiry skills test. The student can only answer four of the 35 questions (11.43%) correctly. Item Q20 is the most difficult for students to answer, while item Q12 is the easiest.

DIF analysis based on university categories is conducted to evaluate whether inquiry skills test item bias is detected between students from public and private universities. Even though, theoretically, students' inquiry skills are not influenced by the university category. This data analysis presentation can complement the analysis of students' inquiry skills using DIF analysis. The results of the DIF analysis based on university categories are shown in Figure 3. DIF analysis based on gender (Figure 4) and semester level (Figure 5) is carried out to evaluate whether any inquiry skills test item bias is detected between males and females in semesters II, IV, and VI. The extreme values for each item are marked with green dotted circle lines in both Figures.

The results of DIF analysis based on gender reveal that three of the 35 items have significant Rasch-Welch test probability ($p<0.05$): Q9 ($p=0.0056$), Q15 ($p= 0.0483$), and Q33 ($p=0.0458$). Items Q9, Q15, and Q33 have DIF contrast of 1.54, 0.91, and 1.04 respectively. Only items Q9 and Q33 have significant DIF and DIF contrast > 1.00. From Figure 3, the difference in peak logit values for items Q9 and Q33 is much higher than that for item Q15. This analysis shows that only two (Q9 and Q33) of the 35 items (5.71%) inquiry skills test questions contain gender-based bias.

The results of DIF analysis based on semester level reveal that 9 of the 35 items have significant Rasch-Welch test probability ($p<0.05$): Q5 ($p=0.0073$), Q9 ($p= 0004$), Q10 ($p=0.0343$), Q15 ($p=0.0103$), Q25 ($p=0.0237$), Q27 ($p=0.0495$), Q30 (p=0.0022), Q34 (p=0.0077), and Q35 (p=0.0158). Items Q5, Q9, Q10, Q15, Q25, Q27, Q30, Q34, and Q35 have DIF contrast of 1.47, 1.42, 1.12, 1.35, 1.28, 0.86, 1.21, 1.20 and 4.69, respectively. Eight of the nine items have a DIF contrast of more than 1.00, and item Q35 has the highest DIF contrast. From Figure 5, the Q35 difference has the highest peak difference. This analysis shows that 8 of the 35 items (22.86%) of inquiry skills test questions contain bias based on semester level.

An inferential statistical test with non-parametric statistics determines differences in students' inquiry skills test scores based on university category, gender, and semester level. Difference analysis inquiry skills based on university category and gender is carried out using two independent sample tests, Mann-Whitney U Test, while differences based on semester level are carried out using the Kruskal Wallis Test. The results of this difference test analysis are displayed in Table 7.

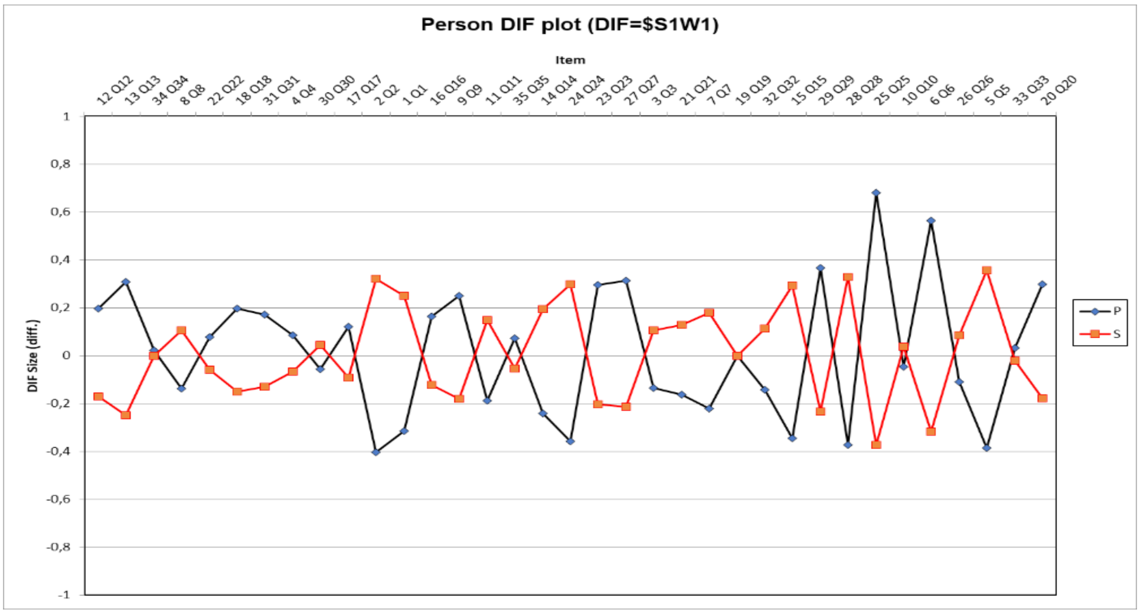

Figure 3. DIF analysis based on the university category (P=Private, S=State)

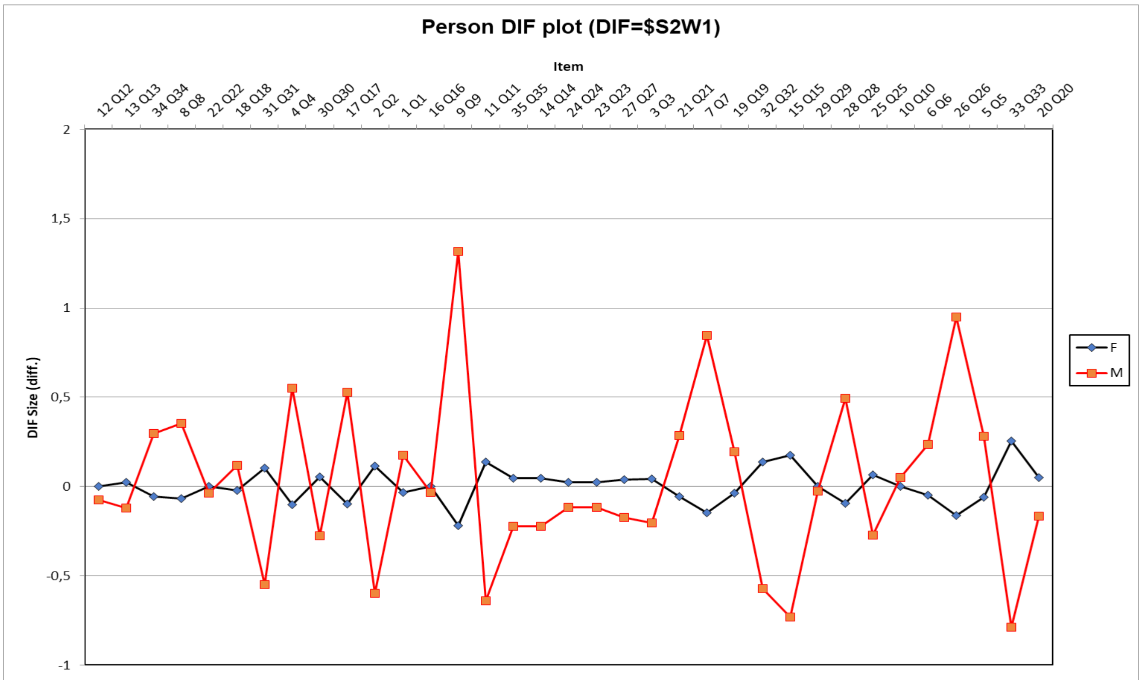

Figure 4. DIF analysis based on the gender (F=Female, M=Male)

The results of non-parametric statistical tests show no significant difference (sig. >0.05) in the inquiry skills of prospective physics teachers based on the university category. In other words, prospective physics teachers from state universities have inquiry skills that are relatively the same as prospective physics teachers from private universities. The same findings can be seen from the differences in students' inquiry skills based on semester level. It shows that after the pandemic, the inquiry skills of students in semesters II, IV, and VI are relatively the same. However, based on gender, there are significant differences in inquiry skills between male and female students. Furthermore, concerning the average inquiry skills score, it is known that the average inquiry skills score for male students is higher than for female students.

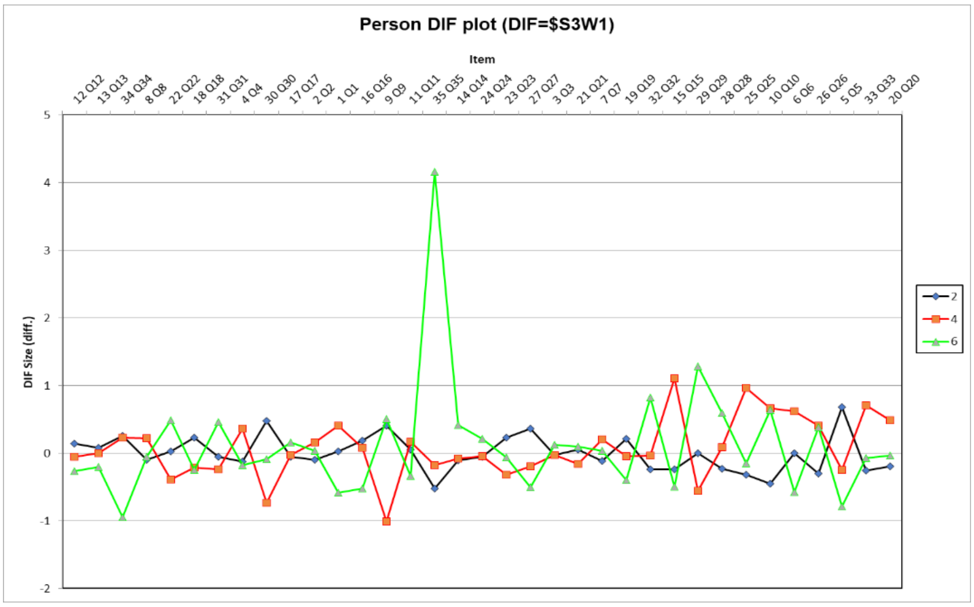

Figure 5. DIF analysis based on the level of semester (2=level II, 4=level IV, 6=level IV)

Table 7. Comparison of inquiry skills based on demographic data

Data	Non-parametric statistic of inquiry skills test	
	Sig.	**Decision**
University Category	0.150	not sig. different
Gender	0.025	Sig. different
Level of Semester	0.189	not sig. different

The results show that the inquiry skills test adapted to evaluate the inquiry skills of prospective physics teachers in Indonesia has met the valid and reliable criteria according to the Rasch Model parameters. The investigation results into the quality of inquiry skills test items show that the range of Rasch parameters based on MNSQ ranges from 0.85 to 1.28, thus meeting the suitability criteria and standard error < 0.5, which confirms the criteria for precise items in measurement. The implication for measurement is that the data has reasonable predictability (Linacre, 2023), so this adapted inquiry skills test can be used well to predict and evaluate the inquiry skills of prospective physics teachers and can be applied to other relevant respondents. The item's separation index and high reliability will give researchers confidence to replicate the placement of inquiry skills test items on other appropriate student respondents (Laliyo et al., 2022).

The mean logit person measure is much lower (negative value) than the value of 0.0 logit (mean item), thus showing a tendency for students to have lower abilities to answer inquiry skills test questions. It is visible from the students' inquiry skills scores, which do not reach 50% of the maximum test score. These low student inquiry skills indicate that during the COVID-19 pandemic, physics learning has not optimally accommodated the inquiry skills of prospective physics teachers. The COVID-19 pandemic has had a significant impact and forced us to rethink the right way to teach and study physics (Monteiro et al., 2023) at that time, including inquiry into physics learning. Although some researchers such as Coramik and Inanç (2023) and Monteiro et al. (2023) have designed simple and flexible physics experimental tools that can be used for inquiry at home, students' inquiry skills have not been trained perfectly.

Person reliability only produces an index in the low category, which means that the respondent's abilities are not diverse enough. It also shows respondents' low ability to replicate the ability to answer the same test items and measure the same constructs (Bond et al., 2021). Person diagnostics from respondents who were diagnosed with Misfit (16SF2 and 52SF2) indicate that neither answered a single easy category item correctly. However, both respondents can correctly answer several items in the difficult category (unexpected success). It can be indicated by guessing, cheating, carelessness, fumbling, plodding, and cultural bias in answering questions (Linacre, 2023; Susongko et al., 2019).

Based on the difficulty level of the items, items Q12 and Q13 are the easiest for students to answer. Item Q12 is an item that represents the skill aspect of formulating a hypothesis, while item Q13 is an item that represents the skill aspect of identifying problems. These two skills are essential as guides in conducting investigations (Tanak, 2020). On the other hand, items Q20 and Q33 are difficult for students. These two items refer to the aspect of predicting skills in inquiry skills. The result is different from □ipková and Karol□ík (2018), which show that students' inquiry skills are lower in planning investigations or experiments. However, these results are strengthened by Saputra et al. (2019), that the predicting skills of prospective physics teachers are still relatively low. This aspect becomes difficult because students think about possible causes before stating the results (Arnold et al., 2018). It implies that the learning strategy implemented should emphasize better predicting skills. So far, the inquiry learning model has positively impacted students' inquiry skills (Firman et al., 2019). However, the inquiry learning model needs to be developed and modified to stimulate students to think critically in formulating better predictions.

DIF analysis shows that the inquiry skills test only contains items with a minimal bias towards differences in university categories (2.86%) and gender (5.71%). Meanwhile, based on semester level, inquiry skills test items contain a bias of 22.86%. Although differences in public and private university financing affect learning facilities and flexibility in curriculum development (Jiawen Wang et al., 2020), the results of this research show different things regarding student inquiry skills. These results confirm that prospective physics teachers from state and private universities have almost the same abilities in inquiry skills. The results of the statistical tests confirm that there are no significant differences in student inquiry skills based on the university category. Student inquiry skills based on semester level also do not differ significantly, although there is bias in the test items. It is similar to Ješková et al. (2018) that there is no difference in students' inquiry skills in the first and second years.

DIF analysis based on gender shows that the inquiry skills test items provide a minimal bias (5.71%). This percentage is also slightly smaller than Susongko et al. (2021) (7%) and Mubarokah et al. (2021) in evaluating students' scientific literacy skills. Differences in student abilities cause DIF based on gender in the inquiry skills test in this study according to the construct of the inquiry skills test, so it is not a test bias. Other studies also find no gender-based bias in science motivation item tests (Rachmatullah et al., 2017) and attitudes toward science (Aini et al., 2019; Huang et al., 2020). However, based on statistical analysis, there are significant differences in the inquiry skills of male and female students. If we look at the average inquiry skills score, it is known that the average inquiry skills score for male students is higher than for female students. These results are similar to Ješková et al. (2018), Nicol et al. (2022), and Jingying Wang et al. (2015), that males' inquiry skills are better than females'. However, other factors still influence students' inquiry skills, such as motivation, attitude, independence, and cognitive skills in learning (Elisanti, 2020), so further research can consider these factors to evaluate the inquiry skills of prospective physics teachers comprehensively.

Since inquiry skills are fundamental for prospective physics teachers, lecturers must develop students' inquiry skills through various learning strategies. Learning strategies during the pandemic, generally carried out online, do not seem effective enough in training students' inquiry skills. After the pandemic ends and adapts to the development of 21st-century education, learning in universities for prospective physics teachers needs to rethink inquiry learning strategies to encourage the acquisition of global competencies, including knowledge, skills, and dispositions (Wu, 2023) and also include comprehensive inquiry skills. Project inquiry learning needs to be interdisciplinary and multidisciplinary (Ozturk, 2021; Santos et al., 2023) in the education curriculum for prospective physics teachers to enable students to develop inquiry competencies globally. Then, developing inquiry skills must also be complemented by carrying out appropriate assessments in the learning process.

CONCLUSION

In conclusion, the results of this research provide an understanding of the inquiry skills of prospective physics teachers. The adapted inquiry skills test instrument has met the validity, reliability, and slight bias criteria based on university category and gender. Therefore, this instrument can be adapted to measure student inquiry skills in other fields of science education. A small number of test items need to be corrected so they are not biased based on the student's semester level. The evaluation results show that male students show better inquiry skills than female students. However, the average inquiry skills of prospective physics teachers are not optimal. Reorganization is needed in planning and implementing learning strategies that can facilitate the inquiry skills of prospective physics teachers.

This research has three weaknesses. First, the respondents in this research are small and focused on respondents in higher education in the education sector on several large islands in Indonesia. So, this research cannot be generalized. Further research can be carried out with a larger sample size in all provinces in Indonesia. Second, the research identifies a separation index of person and person reliability, which is still low, so increasing the range of assessment scales for each inquiry skills test item and designing the test time to be free enough to answer test questions can be considered in further research.

ACKNOWLEDGEMENTS

The authors would like to thank DRTPM (Direktorat Riset, Teknologi dan Pengabdian kepada Masyarakat) Kemdikbudristek-RI for funding the postgraduate research grant in 2023/2024 and also BPPT (Balai Pembiayaan Pendidikan Tinggi) Kemdikbudristek-RI and LPDP (Lembaga Pengelola Dana Pendidikan) Kemenkeu-RI as sponsors of the publication of this article. The authors also thank Mr. Carl J. Wenning for helpful discussions and suggestions in this research.

REFERENCES

Adams, D., Chuah, K. M., Mohamed, A., & Sumintono, B. (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.

Aini, R. Q., Rachmatullah, A., & Ha, M. (2019). Indonesian primary school and middle school students' attitudes toward science: Focus on gender and academic level. *Journal of Baltic Science Education*, *18*(5), 654–667.

Altbach, P. G., & De Wit, H. (2020). *Are we at a transformative moment for online learning?* University World News. https://www.universityworldnews.com/post.php?story=20200427120502132

Andrich, D. (2018). Advances in social measurement: A Rasch measurement theory. In F. Guillemin, A. Leplège, S. Briançon, E. Spitz, & J. Coste (Eds.), *Perceived Health and Adaptation in Chronic Disease* (1st ed., pp. 66–91). Routledge/Taylor & Francis Group.

Arnold, J. C., Boone, W. J., Kremer, K., & Mayer, J. (2018). Assessment of competencies in scientific inquiry through the application of rasch measurement techniques. *Education Sciences*, *8*(4).

Azizan, N. H., Mahmud, Z., Rambli, A., & Hafizah Azizan, N. (2020). Rasch Rating Scale Item Estimates using Maximum Likelihood Approach: Effects of Sample Size on the Accuracy and Bias of the Estimates. *International Journal of Advanced Science and Technology*, *29*(4s), 2526–2531. https://www.researchgate.net/publication/343152554

Bahasoan, A., Ayuandiani, W., & Mukhram, M. (2020). Effectiveness of Online Learning in Pandemic Covid-19. *International Journal of Science, Technology & Management*, *1*(2), 100–106.

Bond, T. G., Yan, Z., & Heene, M. (2021). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (4th ed.). Routledge.

Boone, W. J., & Staver, J. R. (2020). *Advances in Rasch Analyses in the Human Sciences*. Springer.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences* (1st ed.). Springer.

Candra Wicaksono, A. G., & Korom, E. (2023). Attitudes towards science in higher education: Validation of questionnaire among science teacher candidates and engineering students in Indonesia. *Heliyon*, e20023.

Čipková, E., & Karolčík, Š. (2018). Assessing of Scientific Inquiry Skills Achieved by Future Biology Teachers. *Chemistry-Didactics-Ecology-Metrology*, *23*(1–2), 71–80.

Connor, G. O., & Rosicka, C. (2020). Science inquiry skills. *The Australian Council for Educational Research*, 8.

Constantinou, C. P., Tsivitanidou, O. E., & Rybska, E. (2018). What Is Inquiry-Based Science Teaching and Learning? *Contributions from Science Education Research*, *5*, 1–23.

Coramik, M., & Inanç, B. (2023). A physical pendulum experiment with Lego, Phyphox and Tracker. *Physics Education*, *58*(055014), 1–8.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational Research Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (- (ed.); 6th ed.). Pearson Education, Inc.

Egara, F. O. (2023). Gender difference in secondary school students' retention in algebra: A computer simulation approach. *EURASIA Journal of Mathematics, Science and Technology Education*, *19*(7).

Elisanti, E. (2020). Analysis of Students Inquiry Skills in Senior High School Though Learning Based on the Hierarchy of Inquiry Model. *Advances in Social Science, Education and Humanities Research*, *422*(ICOPE 2019), 409–414.

Ezra, O., Cohen, A., Bronshtein, A., Gabbay, H., & Baruth, O. (2021). Equity factors during the COVID - 19 pandemic: Difficulties in emergency remote teaching (ert) through online learning. *Education and Information Technologies*, 7657–7681.

Fackler, A. K., & Sexton, C. M. (2020). Science Teacher Education in the Time of COVID-19: A Document Analysis. *Electronic Journal for Research in Science & Mathematics Education*, *24*(3), 5–13.

Fan, C., Chang, K., Lee, K., Yang, W., & Pakpour, A. H. (2022). Rasch Modeling and Differential Item Functioning of the Self-Stigma Scale-Short Version among People with Three Different Psychiatric Disorders. *International Journal of Environmental Research and Publich Health*, *19*(8843), 1–15.

Fauth, B., Decristan, J., Decker, A. T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education*, *86*, 102882.

Firman, M. A., Ertikanto, C., & Abdurrahman, A. (2019). Description of meta-analysis of inquiry-based learning of science in improving students' inquiry skills. *Journal of Physics: Conference Series*, *1157*(2).

Gao, Z., Zhang, L., Ma, J., Sun, H., Hu, M., Wang, M., Liu, H., & Guo, L. (2023). Reliability and validity of the Chinese version of hte self-directed learning instrument in Chinese nursing students. *BMC Nursing*, *22*(51), 1–9.

Goss-Sampson, M. A. (2020). *Statistical Analysis in JASP: A Guide for Students (JASP v 0.140* (4th ed.). JASP.

Graham, S. R., & Tolar, A. (2020). Teaching Preservice Teachers about COVID-19 through Distance Learning. *Electronic Journal for Research in Science & Mathematics Education*, *24*(3), 29–37.

Hoffmann, L. (2002). Promoting girls' interest and achievement in physics classes for beginners. *Learning and Instruction*, *12*(2002), 447–465.

Huang, F., Huang, L., & Oon, P.-T. (2020). Constructs Evaluation of Student Attitudes Toward Science—A Rasch Analysis. In *Rasch Measurement* (pp. 139–157). Springer Singapore.

JASP Team. (2023). *JASP Software (Version 0.17.2)* (0.17.2.1). University of Amsterdam. https://jasp-stats.org/

Ješková, Z., Balogová, B., & Kireš, M. (2018). Assessing inquiry skills of upper secondary school students. *Journal of Physics: Conference Series*, *1076*(1).

Khalaf, M. A., Mohammed, E., & Omara, N. (2022). Rasch analysis and differential item functioning of English language anxiety scale (ELAS) across sex in Egyptian context. *BMC Psychology*, 1–16.

King, B. M., Rosopa, P., & Minium, E. W. (2018). *Statistical reasoning in the behavioral sciences* (7th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Koh, J. H. L., & Daniel, B. K. (2022). Shifting online during COVID-19: A systematic review of teaching and learning strategies and their outcomes. *International Journal of Educational Technology in Higher Education*, *19*(56), 1–23.

Kuo, C. Y., Wu, H. K., Jen, T. H., & Hsu, Y. S. (2015). Development and Validation of a Multimedia-based Assessment of Scientific Inquiry Abilities. *International Journal of Science Education*, *37*(14), 2326–2357.

Laliyo, L. A. R., Sumintono, B., & Panigoro, C. (2022). Measuring changes in hydrolysis concept of students taught by inquiry model: stacking and racking analysis techniques in Rasch model. *Heliyon*, *8*(3), e09126.

Lamprianou, I. (2020). *Applying the Rasch Model in Social Sciences Using R and BlueSky Statistics* (1st ed.). Routledge/Taylor & Francis Group.

Linacre, J. M. (2023). *A User's Guide to WINSTEPS MINISTEP: Rasch-Model Computer Programs (Program manual 5.6.0).* Winsteps®. https://www.winsteps.com/tutorials.htm

Liu, X. (2020). *Using and Developing Measurement Instruments in Science Education: A Rasch Modeling Approach* (C. S. Kalman (ed.); 2nd ed.). Information Age Publishig, Inc.

Lou, Y., Blanchard, P., & Kennedy, E. (2015). Development and validation of a science inquiry skills assessment. *Journal of Geoscience Education*, *63*(1), 73–75.

Lu, K., Pang, F., & Shadiev, R. (2021). Understanding the mediating effect of learning approach between learning factors and higher order thinking skills in collaborative inquiry-based learning. *Educational Technology Research and Development*, *69*(5), 2475–2492.

Monteiro, M., Stari, C., & Marti, A. C. (2023). A home-lab experiment: resonance and sound speed using telescopic vacuum cleaner pipes. *Physics Education*, *58*(013003), 1–4.

Morrell, P. D., Rogers, M. A. P., Pyle, E. J., Roehrig, G., Veal, W. R., Morrell, P. D., Rogers, M. A. P., Pyle, E. J., & Roehrig, G. (2020). Preparing Teachers of Science for 2020 and Beyond: Highlighting Changes to the NSTA / ASTE Standards for Science Teacher Preparation. *Journal of Science Teacher Education*, *31*(1), 1–7.

Mubarokah, N., Permanasari, A., & Eliyawati, E. (2021). A case study of scientific literacy in natural science subject using rasch analysis model

(RAM). *Journal of Physics: Conference Series*, *1806*(1), 012144.

Murphy, E. (2009). Online synchronous communication in the second-language classroom. *Canadian Journal of Learning and Technology*, *35*(3), 10–22.

National Research Council. (2014). *Developing Assessments for the Next Generation Science Standards* (J. W. Pellegrino, M. R. Wilson, J. A. Koenig, & A. S. Beatty (eds.)). The National Academies Press.

Nicol, C. B., Gakuba, E., & Habinshuti, G. (2022). Students' perceived science inquiry process skills in relation to school type and gender. *Perspectives in Education*, *40*(2), 159–174.

Novitra, F., Fastiyed, Yohandri, & Asrizal. (2021). Development of Online-based Inquiry Learning Model to Improve 21st-Century Skills of Physics Students in Senior High School. *EURASIA Journal of Mathematics, Science and Technology Education*, *17*(22).

Özer, F., & Sarıbaş, D. (2022). Exploring Pre-service Science Teachers' Understanding of Scientific Inquiry and Scientific Practices Through a Laboratory Course. In *Science and Education* (Issue 2022). Springer Netherlands.

Ozturk, E. (2021). The Effect of STEM Activities on the Scientific Inquiry Skills of Pre-service Primary School Teachers. *Journal of Education in Science, Environment and Health*, *7*(4).

Purwati, R., Liestari, S. P., Suwandi, T., Wulan, A. R., & Utari, S. (2021). Profile of Learning Experiences and Students' Scientific Inquiry Skills in Science Subjects. *Proceedings of the International Conference on Educational Assessment and Policy (ICEAP 2020)*, *545*(Iceap 2020), 22–28.

Rachmatullah, A., Octavianda, R. P., Ha, M., Rustaman, N. Y., & Diana, S. (2017). Construct Validity of Science Motivation and Beliefs Instrument (SLA-MB): A Case study in Sumedang, Indonesia. *Journal of Physics: Conference Series*, *812*, 012009.

Raof, S. A., Mustaâ, A. H., Zamzuri, F. K., & Salleh, M. H. (2021). Validity and reliability of students perceptions on OBE approach in Malaysian VC using Rasch model. *Journal of Innovation in Educational and Cultural Research,* 2(2), 44-50.

Sagala, R., Umam, R., Tahir, A., Saregar, A., & Wardani, I. (2019). The Effectiveness of STEM-Based on Gender Differences: The Impact of Physics Concept Understanding. *European Journal of Educational Research*, *8*(3), 753–761.

Santos, C., Rybska, E., Klichowski, M., Jankowiak, B., Jaskulska, S., Domingues, N., Carvalho, D., Rocha, T., Paredes, H., Martins, P., & Rocha, J. (2023). Science education through project-based learning: a case study. *Procedia Computer Science*, *219*(2022), 1713–1720.

Saputra, H., Suhandi, A., & Setiawan, A. (2019). Profile of inquiry skills pre-service physics teacher in Aceh. *Journal of Physics: Conference Series*, *1157*(3).

Setiono, S., Rustaman, N. Y., Rahmat, A., & Anggraeni, S. (2019). Student's inquiry skills and learning achievement in plant anatomy practical work using open-guided inquiry. *Journal of Physics: Conference Series*, *1157*(2). https://doi.org/10.1088/1742-6596/1157/2/022089

Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon*, *7*(11).

Soeharto, S., & Csapó, B. (2022). Assessing Indonesian student inductive reasoning: Rasch analysis. *Thinking Skills and Creativity*, *46*(November 2021).

Spronken-Smith, R., & Walker, R. (2010). Can inquiry-based learning strengthen the links between teaching and disciplinary research? *Studies in Higher Education*, *35*(6), 723–740.

Sunjaya, D. K., Sumintono, B., Gunawan, E., Herawati, D. M. D., & Hidayat, T. (2022). Online Mental Health Survey for Addressing Psychosocial Condition During the COVID-19 Pandemic in Indonesia: Instrument Evaluation. *Psychology Research and Behavior Management*, *15*, 161–170.

Surahman, E., & Wang, T. H. (2023). In-service STEM teachers' professional development programmes: A systematic literature review 2018–2022. *Teaching and Teacher Education*, *135*(July), 104326.

Susongko, P., Arfiani, Y., & Kusuma, M. (2021). Determination of Gender Differential Item Functioning in Tegal-Students' Scientific Literacy Skills with Integrated Science (SLiSIS) Test Using Rasch Model. *Jurnal Pendidikan IPA Indonesia*, *10*(2), 270–281.

Susongko, P., Kusuma, M., & Widiatmo, H. (2019). Using Rasch Model to Detect Differential Person Functioning and Cheating Behavior in Natural Sciences Learning Achievement Test. *Jurnal Penelitian Dan Pembelajaran IPA*, *5*(2), 94.

Syawaludin, A., Prasetyo, Z. K., Safruddin, C., & Jabar, A. (2022). The Effect of Project-based Learning Model and Online Learning Settings on Analytical Skills of Discovery Learning, Interactive Demonstrations, and Inquiry Lessons. *Journal of Turkish Science Education*, *19*(2), 608–621.

Tanak, A. (2020). Developing Pre-service Science Teachers' Teaching Practices with an Emphasis on Higher Order Thinking. *Science Education International*, *31*(3), 237–246.

Teig, N., & Nilsen, T. (2022). Profiles of instructional quality in primary and secondary education: Patterns, predictors, and relations to student achievement and motivation in science. *Studies in Educational Evaluation*, *74*(September 2020), 101170.

Wang, Jiawen, Yang, M., & Maresova, P. (2020). Sustainable development at higher education in China: A comparative study of students' perception in public and private universities. *Sus-*

tainability (Switzerland), *12*(6).

Wang, Jingying, Guo, D., & Jou, M. (2015). A study on the effects of model-based inquiry pedagogy on students' inquiry skills in a virtual physics lab. *Computers in Human Behavior*, *49*(2015), 658–669.

Wenning, C. J. (2007). Assessing inquiry skills as a component of scientific literacy. *Journal of Physics Teacher Education Online*, *4*(2), 21–24.

Wenning, C. J. (2010). Levels of inquiry: Using inquiry spectrum learning sequences to teach science. *Journal of Physics Teacher Education Online*, *5*(3), 11–20.

Wenning, C. J., & Vieyra, R. E. (2020). Teaching High School Physics: Interacting with Physics Students. In *AAPT Physics Education* (1st ed.). AIP Publishing LLC. https://doi.org/10.1063/9780735422056

Wu, X. (2023). Fostering Chinese preservice teachers' global competence through inquiry learning in glocalised educational contexts. *Teaching and Teacher Education*, *122*.