Migration Letters

Volume: 21, No: 4, pp. 173-186

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)
www.migrationletters.com

Bridging Gaps in Classrooms: Exploring the Intersection between Migration Realities and Differentiated Instruction Using Diagnostic Test

Slamet Maulana[1], Ani Rusilowati[2], Sunyoto Eko Nugroho[3], Endang Susilaningsih[4]

Abstract

The gap in students' basic knowledge is often the primary problem when starting the learning process in the classroom. The phenomenon of migration is one of the various factors causing gaps in the school. Mutation students tend to have high basic knowledge. Differentiated instruction is a solution to dealing with student diversity in the classroom. This study aims to develop diagnostic test instruments to realize differentiated instruction that can bridge the gap in fundamental knowledge in the school due to migration. This research uses research and development design with the ADDIE model. Based on the data collected, there appears to be a gap between local and migration students. In addition, the data collection results also show that the achievement of learning completeness by mutation students is higher than local students. This research produces diagnostic test instruments to realize differentiated instruction that can bridge gaps in the classroom due to migration. The results showed that students were divided into three categories. This classification is then used as reference to provide learning design recommendations for each category. These recommendations aim to assist teachers in developing future lesson plans and realizing differentiated instruction in the classroom. At the end of this research, it has succeeded in developing diagnostic test instruments for differentiated instruction to bridge the gap in fundamental knowledge in the classroom due to migration. The differentiated instruction in the school has succeeded in creating equality in the learning process, especially in the realm of knowledge.

Keywords: *Migration Realities, Differentiated Instruction, Diagnostic Test.*

Introduction

Gaps in a classroom are still an educational problem. Gaps in a class take many forms. The gap in students' basic knowledge is often the primary problem when starting the learning process in the classroom.

The gap phenomenon will lead to new problems. Students with low basic knowledge compared to peers in the classroom tend to have low motivation (Kopeyev et al., 2020). In addition, gaps are an essential factor that causes failures in applying learning analysis

[1] Ph.D. student, Science Education Program Study, Faculty of Mathematics and Natural Science, Universitas Negeri Semarang, Semarang, Central Java, Indonesia
[2] Adjunct Professor, Science Education Program Study, Faculty of Mathematics and Natural Science, Universitas Negeri Semarang, Semarang, Central Java, Indonesia
[3] Associate Professor, Physics Education Program Study, Faculty of Mathematics and Natural Science, Universitas Negeri Semarang, Semarang, Central Java, Indonesia
[4] Associate Professor, Chemistry Education Program Study, Faculty of Mathematics and Natural Science, Universitas Negeri Semarang, Semarang, Central Java, Indonesia

in the classroom (Prieto et al., 2019). Several factors can cause gaps in a class. The phenomenon of migration is one of the various factors causing gaps in the classroom.

Migration is the process of moving people from one location to another. This migration phenomenon is carried out to disseminate and equalize. In addition, the phenomenon of migration is usually carried out by the community to continue its survival due to economy, politics, climatic factors, availability of foodstuffs, and so on (Nienkerke et al., 2023; Smith & Wesselbaum, 2020; Urbański, 2022; Vinke et al., 2020). As a result, people can live better, economic distribution is more even, and many other benefits from this migration phenomenon

Behind this good goal, the phenomenon of migration raises gaps among students. Based on observations, students from cities tend to have high basic knowledge. Conversely, students who come from rural areas tend to have low knowledge. This creates gaps for teachers in carrying out the learning process in the classroom.

Differentiated instruction is a solution to dealing with student diversity in the classroom. The application of differentiated instruction in the classroom reduces the gap between learners. This differentiated instruction can also increase motivation, improve relationships between students and teachers, increase scientific process skills development, increase scientific literacy levels, and narrow achievement gaps among students (Ginja & Chen, 2020; Şentürk & Sari, 2018). As the achievement gap among students shrinks, diversity in the classroom is reduced (Magableh & Abdullah, 2020).

Differentiated instruction provides teachers with a starting point to meet the learning needs of diverse learners in creating equal learning opportunities (Alstete et al., 2020; Kamarulzaman et al., 2021; Maulana et al., 2020; Ramli & Nurahimah, 2020; Smale-Jacobse et al., 2019; Whitley et al., 2021). Teachers can use differentiated instruction to improve learning skills and contribute to the student’s cognitive and socio-emotional development (Zafiri et al., 2019). In addition, improvements in the quality of assessment and the level of active participation of learners occur due to the implementation of differentiated instruction (Palieraki & Koutrouba, 2021).

However, the implementation of differentiated instruction is still limited. The observations show that teachers still do not understand the concept and essence of differentiation learning. In addition, the results of observations also show that teachers still do not have special instruments that can map students' basic abilities.

Differentiated instruction is an essential but complex teaching skill many teachers have not mastered and feel unprepared for (Onyishi & Sefotho, 2020; van Geel et al., 2019). Lack of resources, time, support, teacher knowledge and competence, class size, and learning assessment hinder differentiated instruction in the classroom (Shareefa, 2021; Shareefa et al., 2019). Therefore, due to this migration phenomenon, a particular instrument is needed to realize differentiated instruction in the classroom.

The diagnostic test instrument becomes a particular instrument used to map the basic knowledge of each student. Through diagnostic test instruments, teachers will be facilitated in classifying students based on their basic abilities. Therefore, developing diagnostic test instruments is one of the efforts to realize differentiated instruction and bridge the gap due to migration.

Nevertheless, diagnostic test instruments will be developed to map students' basic abilities and provide learning design recommendations for each student. A good diagnostic test can provide an accurate picture of the misconceptions a student is experiencing based on the information he or she made mistakes (Wilantika et al., 2018). Therefore, differentiated instruction can be easier to realize. As a result, the gap in basic knowledge between students due to migration can be resolved.

Based on these descriptions, this study aims to develop diagnostic test instruments to realize differentiated instruction that can bridge the gap in fundamental knowledge in the classroom due to migration.

Method

This research uses research and development design with the ADDIE model. The ADDIE model was chosen because of its simple and easy steps to apply in this research process. The ADDIE model consists of five stages: Analyze, Design, Develop, implement, and Evaluate.

At the analysis stage, research is carried out by conducting preliminary studies. This preliminary study consists of a literature study and a field study. This preliminary study was carried out for initial data collection related to the phenomenon of migration, problems that teachers have in the classroom, and the needs of teachers to realize a maximum learning process in the classroom.

At the design stage, research is carried out by developing a grid of question items. In addition, at this stage, a blueprint of the diagnostic test instrument to be developed is also made.

At the development stage, research is done by making the expected diagnostic test instruments. The diagnostic test instruments that have been developed are then validated theoretically and empirically. Theoretical validation was carried out by involving ten experts through the calculation of CVR (Content Validity Ratio) and Aikens' V values for validation of the content of each question item, as well as CVI (Content Validity Index) and questionnaires for theoretical validation of diagnostic test instruments. The empirical validation is carried out by involving a minimum of 200 students to try to answer each question on the diagnostic test instrument that has been developed. The results of students' answers are then analyzed using the Rasch Model to determine their quality. The qualities tested in this case are construct validity, reliability, item fit, difficulty level, and the bias of the question grain.

At the implementation stage, research is carried out by piloting diagnostic test instruments that are developed and validated. The trial was conducted with a minimum of 200 students. The trial begins with mapping students' initial knowledge. After that, the teacher will carry out differentiated learning based on the category of each student. After carrying out the lesson, students will carry out formative tests. Formative test results are then compared between local students and mutations, then analyzed.

At the end of each stage, the evaluation process is carried out. The final evaluation is carried out to obtain the conclusion of the research that has been done.

Result and Discussion

The phenomenon of migration today is still frequent. This phenomenon has a good purpose, which is equity. However, this phenomenon also brings new problems, especially in education. This migration phenomenon results in gaps in fundamental knowledge between one student and another.

This study tries to record the phenomenon of migration in a school in an area of Indonesia (Bekasi). This process involves 261 students. The results of data collection can be seen in Table 1 below.

Table 1. Migration student data.

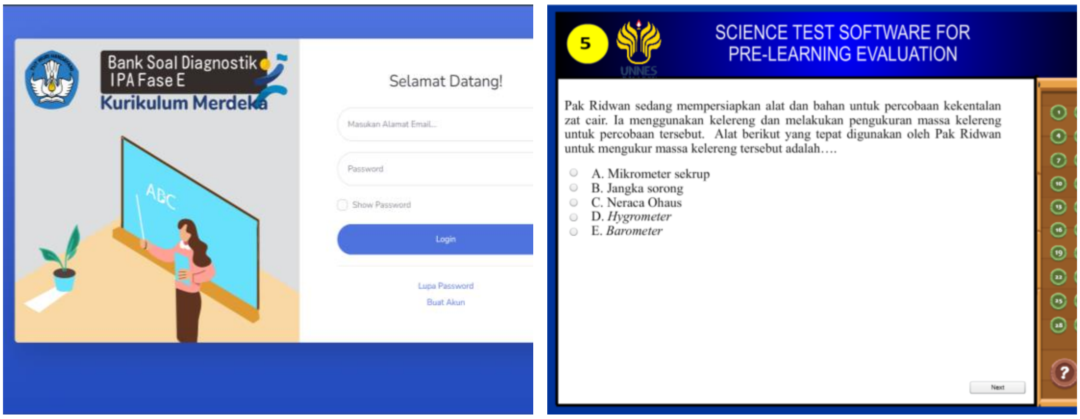
Number	Indicator	Result
1	Student Percentage	Data shows that 198 students (75.86%) are students from the local area, while the remaining 63 students (24.14%) are students from outside the area (students who experienced the migration process).
2	Initial Knowledge Score	Data shows that students from their home regions have an average fundamental knowledge score of 70.84, while, migration students have an average fundamental knowledge score of 78.14, with a maximum score of 99 and a minimum score of 51.
3	Learning Achievement Percentage	The achievement results in previous materials (magnitude and unit), mutation students have a higher percentage of achievement of learning objectives. There are only 74.24% (147 students) from local students that had learning achievements. Meanwhile, 80.95% (51 students) from mutation students that had learning achievements.
4	Migration Origin	The data showed that of the 63 students who carried out the migration process, as many as 27 students (42.86%) were students migrating from the same city but different regions; A total of 20 students (31.75%) were migration students not from the same city, but still in the same province; A total of 7 students (11.11%) are migratory students from other provinces, but still one island; as many as four students (6.35%) are migratory students from other islands in Indonesia; A total of 5 students (7.94%) were students migrating from abroad.
5	Migration Reasons	Data shows that of the 63 students who carried out the migration process, as many as seven students (11.11%) migrated due to natural disasters in their home areas, such as earthquakes, Mount Merapi, and landslides; A total of 31 students (49.21%) migrated due to economic factors; A total of 13 students (20.63%) migrated due to political factors; A total of 3 students (4.76%) migrated due to religious factors; and as many as nine students (14.29%) migrated due to natural conditions such as changes in the contour of the land where they lived.

Based on the data collected above, there appears to be a gap between local and migration students. The focus of the gap in this study lies in the average fundamental knowledge score. There is an average score difference of more than 7 points. Maybe at first glance, the difference in scores is not significant, but the difference shows a gap in basic knowledge between local and migration students.

In addition, the data collection results also show that the achievement of learning completeness by mutation students is higher than those from the original area. It appears that the difference in basic abilities possessed by mutation students compared to original students affects the achievement of student learning completeness.

Following up on migration-related issues, the role of differentiated instruction in the classroom is needed. In order to realize differentiated instruction, further mapping is needed related to the basic knowledge of each student. Therefore, the diagnostic test developed in this study has a considerable role.

This research produces diagnostic test instruments to realize differentiated instruction that can bridge gaps in the classroom due to migration. The focus of the question material on the diagnostic test instrument in this study is measurement in scientific work. The diagnostic test instrument that has been developed has several parts, namely the homepage, token page, test page, and result page.



(a) (b)

Figure 1. (a) Homepage; (b) Test page.

On the home page, students will be asked to enter the username and password their respective teachers gave. After that, the view will switch to the page token. On the token page, students will enter the exam code they will take. The teacher gives the exam code before starting the exam.

After entering the exam code, students will start taking diagnostic tests. On the test page, students can do more straightforward questions by clicking on the question number. After students finish taking the exam, the diagnostic test instrument will display their exam results. Meanwhile, the teacher will obtain the results of student work in his account, along with mapping mastery of basic concepts and learning design recommendations for each student.

The question items developed and used in the diagnostic test instrument are 30 questions. The validity of the content of the question items is done by calculating the CVR and Aiken's V values of each item. Content validity ensures that items and test instruments designed and developed correctly can be measured (Al-Taweel & Awad, 2020). Each question item's content is determined by asking for validity from ten experts. Thus, the minimum CVR value for question items declared valid is 0.62. The minimum score of Aiken's V for question items is declared valid with an error of 5%, 0.70.

The results of calculating the CVR value show that the value of each question item is 0.80 or 1. Each question item is declared valid based on the CVR value. This is in line with previous research on the development of an instrument, which states that the validity of content (through CVR calculations) is essential and crucial (Alizadeh-Siuki et al., 2020; Waltz et al., 2016). The results of calculating Aiken's V value also show that the value of each question item is at least 0.70. This informs that each question item is valid based on Aiken's V score. Based on the results of the two calculations, all question items developed can be declared valid.

After testing the validity of all question items, all valid question items will be combined into a complete diagnostic test instrument. The diagnostic test instrument is then validated through CVI calculations and reviewed by ten experts through questionnaires.

The CVI calculation results show a value of 0.55. The CVI value indicates that the diagnostic test instrument has excellent validity. In addition, a review conducted by ten experts gave theoretical validation results of 81.00% with outstanding criteria. This shows that the diagnostic test instrument has excellent feasibility. The details of reviews from several experts can be seen in Figure 2.

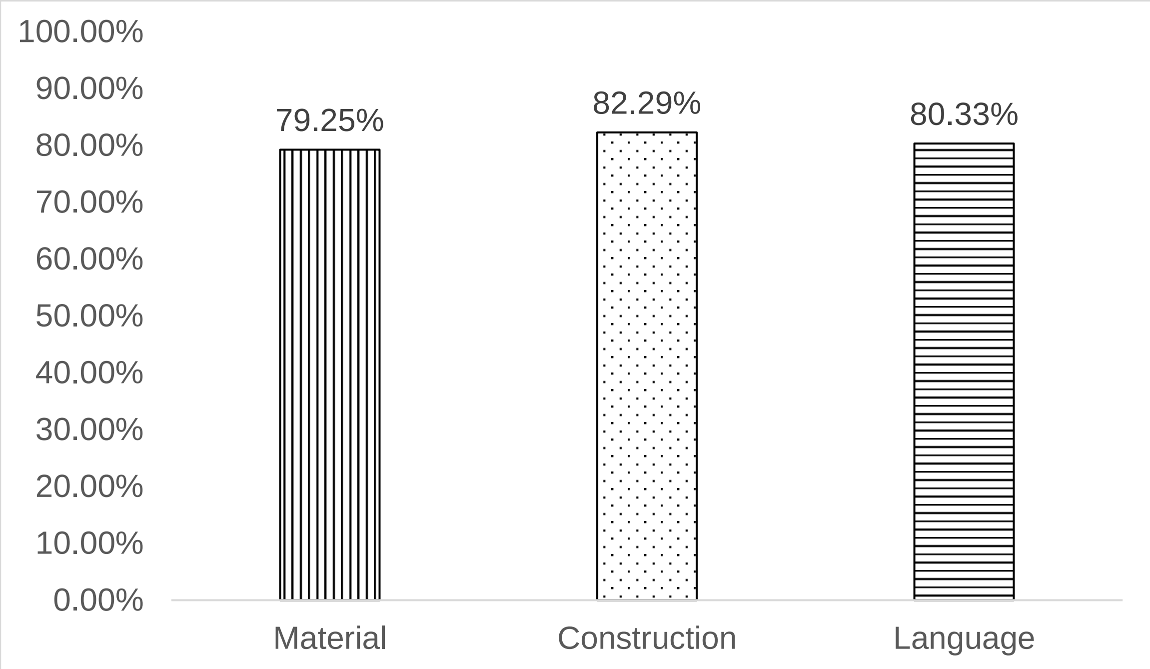


Figure 2. The quality of diagnostic test instruments.

Figure 2 shows that the "material" aspect has good criteria, while the "construction" aspect and the "language" aspect with outstanding criteria. These aspects show that the diagnostic test instrument is suitable for use because it has met the criteria of good and excellent in every aspect.

After theoretical validity is implemented, empirical validity is carried out. Empirical validity in this study involved 311 students. All students are asked to use the diagnostic test instrument. The answer data provided by each student is then analyzed for empirical validity using the Rasch Model.

The Rasch Model was chosen in this study because it can define the validity of instrument constructs (Mohamad et al., 2015). Through the Rasch Model, each question item can be tested again for consistency, where when the question item has been declared fit, it is enough to inform that the related question item is valid and can measure what should be measured (Planinic et al., 2019). The first empirical validity produced is the construct validity of the diagnostic test instruments that have been developed. The results of construct validity analysis by the Rasch Model can be seen in Figure 3.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

		Eigenvalue	Observed		Expected
Total raw variance in observations	=	29.6394	100.0%		100.0%
Raw variance explained by measures	=	9.6394	32.5%		33.5%
Raw variance explained by persons	=	1.2975	4.4%		4.5%
Raw Variance explained by items	=	8.3420	28.1%		29.0%
Raw unexplained variance (total)	=	20.0000	67.5%	100.0%	66.5%
Unexplned variance in 1st contrast	=	11.7774	39.7%	58.9%	
Unexplned variance in 2nd contrast	=	8.2141	27.7%	41.1%	
Unexplned variance in 3rd contrast	=	.0050	.0%	.0%	
Unexplned variance in 4th contrast	=	.0035	.0%	.0%	
Unexplned variance in 5th contrast	=	.0000	.0%	.0%	

Figure 3. Construct validity of diagnostic test instruments.

The results of the Rasch Model analysis showed that the diagnostic test instruments had a Raw Variance value of 33.5%. This informs that the diagnostic test instrument set has good construct validity. This informs that the diagnostic test instrument set actually measures what it should measure

Raw variance is a latent characteristic in a measuring instrument (Al-Zoubi et al., 2018; Nielsen & Dammeyer, 2019) in verifying the construct measurement of an instrument (Chang et al., 2020; Lo et al., 2015). The construct validity explains the content's validity and the consequences of the scoring (Mokshein et al., 2019). Using the Rasch analysis, the research results can establish the item's conformity to the identified construction (Md Yunos et al., 2017) against the test instrument itself in general (Yasin et al., 2015; Yasin et al., 2018).

The second empirical validity produced is the reliability of the diagnostic test instruments. The reliability analysis results by the Rasch Model can be seen in Figure 4.

SUMMARY OF 30 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	11.8	3.0	.00	.88	.96	.1	.97	.1
P.SD	1.3	.0	.92	.12	.60	1.0	.63	.9
S.SD	1.3	.0	.95	.12	.61	1.0	.64	.9
MAX.	14.0	3.0	1.35	1.05	2.35	1.8	2.49	1.8
MIN.	10.0	3.0	-1.53	.77	.16	-2.0	.19	-1.6

REAL RMSE .38 TRUE SD .66 SEPARATION 1.71 Item RELIABILITY .72
MODEL RMSE .38 TRUE SD .66 SEPARATION 1.74 Item RELIABILITY .72
S.E. OF Item MEAN = .21

Figure 4: Reliability of diagnostic test instruments.

The analysis results with the Rasch Model show that the diagnostic test instrument has a reliability value of 0.72. This value indicates that the diagnostic test instrument has good reliability. Reliability provides detailed information regarding the characteristics of the items in the instrument, with the test taker responding to those items (Bodzin et al., 2020). Reliability also informs the consistency of related instruments (Koçak, 2020).

The third empirical validity produced is the determination of the fit of each question item from the diagnostic test instrument. The results of the fit analysis of question items by the Rasch Model can be seen in Figure 5.

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
16	21	30	49.34	4.06	1.13	.8	1.22	1.1	A-.15	.19	66.7	70.1	Q16
19	22	30	47.64	4.20	1.06	.4	1.17	.8	B-.02	.18	73.3	73.3	Q19
23	21	30	49.34	4.06	1.04	.3	1.11	.6	C .06	.19	66.7	70.1	Q23
2	20	30	50.94	3.95	1.05	.4	1.07	.5	D .06	.19	66.7	67.2	Q2
24	22	30	47.64	4.20	1.07	.4	1.07	.4	E .02	.18	73.3	73.3	Q24
28	22	30	47.64	4.20	1.07	.4	1.07	.4	F .03	.18	73.3	73.3	Q28
10	19	30	52.46	3.86	1.04	.4	1.06	.5	G .09	.20	63.3	64.4	Q10
29	23	30	45.80	4.38	1.02	.2	1.06	.3	H .10	.17	76.7	76.7	Q29
1	22	30	47.64	4.20	1.02	.2	1.04	.2	I .12	.18	73.3	73.3	Q1
17	22	30	47.64	4.20	1.04	.3	1.04	.3	J .08	.18	73.3	73.3	Q17
21	21	30	49.34	4.06	1.04	.3	1.04	.3	K .09	.19	66.7	70.1	Q21
5	19	30	52.46	3.86	1.03	.3	1.01	.2	L .14	.20	63.3	64.4	Q5
7	18	30	53.92	3.80	.99	.0	1.03	.3	M .19	.20	66.7	61.9	Q7
13	21	30	49.34	4.06	1.03	.3	1.02	.2	N .12	.19	66.7	70.1	Q13
27	19	30	52.46	3.86	1.02	.2	1.00	.1	O .17	.20	63.3	64.4	Q27
22	20	30	50.94	3.95	1.00	.0	1.01	.1	o .19	.19	66.7	67.2	Q22
26	22	30	47.64	4.20	1.01	.1	.98	.0	n .17	.18	73.3	73.3	Q26
4	20	30	50.94	3.95	1.00	.0	.98	.0	m .21	.19	66.7	67.2	Q4
12	19	30	52.46	3.86	1.00	.0	1.00	.1	l .20	.20	70.0	64.4	Q12
18	23	30	45.80	4.38	.99	.1	.94	-.2	k .22	.17	76.7	76.7	Q18
20	21	30	49.34	4.06	.97	-.1	.95	-.2	j .27	.19	73.3	70.1	Q20
3	21	30	49.34	4.06	.96	-.2	.93	-.3	i .30	.19	73.3	70.1	Q3
8	16	30	56.75	3.74	.96	-.5	.95	-.5	h .30	.20	63.3	58.3	Q8
30	21	30	49.34	4.06	.96	-.2	.92	-.3	g .30	.19	73.3	70.1	Q30
11	19	30	52.46	3.86	.95	-.3	.95	-.3	f .30	.20	70.0	64.4	Q11
15	20	30	50.94	3.95	.94	-.3	.93	-.4	e .32	.19	73.3	67.2	Q15
25	21	30	49.34	4.06	.94	-.3	.94	-.2	d .31	.19	73.3	70.1	Q25
9	19	30	52.46	3.86	.89	-.9	.86	-1.0	c .47	.20	63.3	64.4	Q9
14	21	30	49.34	4.06	.89	-.6	.84	-.8	b .46	.19	73.3	70.1	Q14
6	21	30	49.34	4.06	.87	-.8	.83	-.9	a .50	.19	73.3	70.1	Q6
MEAN	20.5	30.0	50.00	4.03	1.00	.0	1.00	.0			69.9	69.0	
P.SD	1.5	.0	2.41	.16	.06	.4	.09	.5			4.3	4.3	

Figure 6: Item fit diagnostic test instrument.

Each item is declared fit if it meets at least one of the three requirements of the fit item. The three conditions are the means-square outfit value between 0.5 and 1.5, the z-standard outfit value between -2.0 and 2.0, and the point measure correlation value between 0.4 and 0.85. The results of the analysis with the Rasch Model show that each question item in the diagnostic test instrument meets at least two fit item requirements. Therefore, the Rasch Model analysis results show that all question items in the diagnostic test instrument are fit.

The Rasch model is used because it is capable of defining constructs of valid items and providing a clear definition of measurable constructions consistent with theoretical expectations (Mohamad et al., 2015). As for the questions analyzed using the Rasch Model measurement model and declared fit, then the results are enough to provide information that the questions developed are valid and can measure what should be measured (Boone, 2016; Planinic et al., 2019)

The fourth empirical validity produced is the determination of the level of difficulty of each question item from the diagnostic test instrument. The results of the difficulty level analysis of each question item by the Rasch Model can be seen in Figure 6.

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
8	16	30	56.75	3.74	.96	-.5	.95	-.5	.30	.20	63.3	58.3	Q8
7	18	30	53.92	3.80	.99	.0	1.03	.3	.19	.20	66.7	61.9	Q7
5	19	30	52.46	3.86	1.03	.3	1.01	.2	.14	.20	63.3	64.4	Q5
9	19	30	52.46	3.86	.89	-.9	.86	-1.0	.47	.20	63.3	64.4	Q9
10	19	30	52.46	3.86	1.04	.4	1.06	.5	.09	.20	63.3	64.4	Q10
11	19	30	52.46	3.86	.95	-.3	.95	-.3	.30	.20	70.0	64.4	Q11
12	19	30	52.46	3.86	1.00	.0	1.00	.1	.20	.20	70.0	64.4	Q12
27	19	30	52.46	3.86	1.02	.2	1.00	.1	.17	.20	63.3	64.4	Q27
2	20	30	50.94	3.95	1.05	.4	1.07	.5	.06	.19	66.7	67.2	Q2
4	20	30	50.94	3.95	1.00	.0	.98	.0	.21	.19	66.7	67.2	Q4
15	20	30	50.94	3.95	.94	-.3	.93	-.4	.32	.19	73.3	67.2	Q15
22	20	30	50.94	3.95	1.00	.0	1.01	.1	.19	.19	66.7	67.2	Q22
3	21	30	49.34	4.06	.96	-.2	.93	-.3	.30	.19	73.3	70.1	Q3
6	21	30	49.34	4.06	.87	-.8	.83	-.9	.50	.19	73.3	70.1	Q6
13	21	30	49.34	4.06	1.03	.3	1.02	.2	.12	.19	66.7	70.1	Q13
14	21	30	49.34	4.06	.89	-.6	.84	-.8	.46	.19	73.3	70.1	Q14
16	21	30	49.34	4.06	1.13	.8	1.22	1.1	-.15	.19	66.7	70.1	Q16
20	21	30	49.34	4.06	.97	-.1	.95	-.2	.27	.19	73.3	70.1	Q20
21	21	30	49.34	4.06	1.04	.3	1.04	.3	.09	.19	66.7	70.1	Q21
23	21	30	49.34	4.06	1.04	.3	1.11	.6	.06	.19	66.7	70.1	Q23
25	21	30	49.34	4.06	.94	-.3	.94	-.2	.31	.19	73.3	70.1	Q25
30	21	30	49.34	4.06	.96	-.2	.92	-.3	.30	.19	73.3	70.1	Q30
1	22	30	47.64	4.20	1.02	.2	1.04	.2	.12	.18	73.3	73.3	Q1
17	22	30	47.64	4.20	1.04	.3	1.04	.3	.08	.18	73.3	73.3	Q17
19	22	30	47.64	4.20	1.06	.4	1.17	.8	-.02	.18	73.3	73.3	Q19
24	22	30	47.64	4.20	1.07	.4	1.07	.4	.02	.18	73.3	73.3	Q24
26	22	30	47.64	4.20	1.01	.1	.98	.0	.17	.18	73.3	73.3	Q26
28	22	30	47.64	4.20	1.07	.4	1.07	.4	.03	.18	73.3	73.3	Q28
18	23	30	45.80	4.38	.99	.1	.94	-.2	.22	.17	76.7	76.7	Q18
29	23	30	45.80	4.38	1.02	.2	1.06	.3	.10	.17	76.7	76.7	Q29
MEAN	20.5	30.0	50.00	4.03	1.00	.0	1.00	.0			69.9	69.0	
P.SD	1.5	.0	2.41	.16	.06	.4	.09	.5			4.3	4.3	

Figure 6: The level of difficulty of diagnostic test instruments.

The results of the Rasch Model analysis inform that all question points in the instrument are divided into three levels of difficulty: easy, medium, and difficult. The analysis results showed a Mean value of 50.00 and an S.D. value of 2.41. Each question item with a Mean value of more than 52.41 is included in the difficult question category. In contrast, the question item with a Mean value lower than 47.59 is included in the easy question category. The question item with a Mean value between 47.59 and 52.41 is included in the medium question category. The analysis results show that questions 8, 7, 5, 9, 10, 11, 12, and 27 are difficult, questions 18 and 19 are easy, and the rest have moderate difficulty.

The varying difficulty level of the questions is one of the criteria for a suitable measuring instrument (Barus et al., 2019; Hamdu et al., 2020). With variations in the difficulty level of question items in an instrument, the test instrument can measure test takers' ability well. In addition, the difficulty of the question items does not have to be sequential in a test instrument (Isnani et al., 2019).

The fifth empirical validity produced is the determination of the existence of biased question items from the diagnostic test instrument. Analysis of potential bias of question items is divided into two categories, namely bias against gender and bias against school origin. Gender bias is divided into two types: male and female. At the same time, the bias against school origin is divided into three types: students from high schools (general), students from faith-based high schools, and students from vocational high schools. The results of the analysis of the potential bias of each question item by the Rasch Model can be seen in Figure 7.

DIF class/group specification is: DIF=$S4W1

Person CLASSES	SUMMARY DIF CHI-SQUARED	D.F.	PROB.	BETWEEN-CLASS/GROUP UNWTD MNSQ	t=ZSTD
2	1.4196	1	.2335	1.5678	.8144
2	3.9383	1	.0572	4.7545	1.9171
2	.7976	1	.3718	.8537	.3624
2	.9917	1	.3193	1.0709	.5204
2	.2823	1	.5952	.2967	-.2350
2	2.6633	1	.1027	3.0585	1.4293
2	.4707	1	.4927	.4978	.0313
2	2.3176	1	.1279	2.6236	1.2758
2	.0568	1	.8116	.0592	-.8231
2	.2823	1	.5952	.2967	-.2350
2	.0568	1	.8116	.0592	-.8231
2	.0568	1	.8116	.0592	-.8231
2	.0094	1	.9226	.0098	-1.1958
2	2.6633	1	.1027	3.0585	1.4293
2	1.7137	1	.1905	1.8952	.9752
2	.4971	1	.4808	.5269	.0634
2	.1525	1	.6961	.1596	-.4992
2	2.7541	1	.0970	3.3307	1.5181
2	.1525	1	.6961	.1596	-.4992
2	2.1149	1	.1459	2.4049	1.1922
2	.0094	1	.9226	.0098	-1.1958
2	.0512	1	.8210	.0534	-.8512
2	2.6633	1	.1027	3.0585	1.4293
2	1.4196	1	.2335	1.5678	.8144
2	.7976	1	.3718	.8537	.3624
2	1.4196	1	.2335	1.5678	.8144
2	.2823	1	.5952	.2967	-.2350
2	1.4196	1	.2335	1.5678	.8144
2	.8138	1	.3670	.8763	.3800
2	.4971	1	.4808	.5269	.0634

(a)

DIF class/group specification is: DIF=$S3W1

Person CLASSES	SUMMARY DIF CHI-SQUARED	D.F.	PROB.	BETWEEN-CLASS/GROUP UNWTD MNSQ	t=ZSTD
3	1.7486	2	.4136	1.0096	.3429
3	4.6654	2	.0952	3.1089	1.7118
3	.0637	2	.9703	.0340	-1.6952
3	1.4476	2	.4815	.8255	.1476
3	1.5925	2	.4475	.9089	.2393
3	1.9379	2	.3758	1.1839	.5069
3	.0760	2	.9645	.0404	-1.6374
3	2.1799	2	.3326	1.2830	.5932
3	1.4911	2	.4711	.8533	.1788
3	3.0629	2	.2131	1.8764	1.0336
3	1.4911	2	.4711	.8533	.1788
3	.6573	2	.7189	.3584	-.5356
3	3.1422	2	.2048	1.9394	1.0745
3	.5002	2	.7786	.2719	-.7232
3	3.5852	2	.1639	2.4161	1.3589
3	2.6458	2	.2630	1.6733	.8949
3	.3682	2	.8325	.1987	-.9159
3	.1546	2	.9274	.0824	-1.3610
3	.8228	2	.6609	.4524	-.3636
3	.0637	2	.9703	.0340	-1.6952
3	.7682	2	.6795	.4221	-.4163
3	1.4476	2	.4815	.8255	.1476
3	1.1622	2	.5564	.6540	-.0626
3	.1679	2	.9213	.0897	-1.3239
3	1.8748	2	.3880	1.0816	.4128
3	3.7183	2	.1533	2.4407	1.3725
3	.6573	2	.7189	.3584	-.5356
3	1.8050	2	.4019	1.0875	.4184
3	.1546	2	.9274	.0824	-1.3610
3	2.4278	2	.2935	1.5036	.7702

(b)

Figure 1. Bias items based on (a) gender and (b) school origin.

The results of the Rasch Model analysis show that both in terms of gender and school origin, the Prob score on DIF is more than 0.05. This value informs that each question item in the diagnostic test instrument does not experience bias towards gender or school origin. It also confirms that no student from a particular group benefits when working on the questions on the diagnostic test instrument.

The DIF score examines the likelihood of a question item experiencing bias caused by responses or answer patterns from different groups (Rahmani, 2018; Veas et al., 2016; Xu et al., 2020). Biased questions arise because two groups have the same abilities, but there are significant differences when answering question items (Alwi, 2017). This means a particular group will later benefit when answering related question items.

In addition to providing an analysis of the quality of test instruments, the application of Rasch analysis also contains information about the quality of students who take the test (Maulana et al., 2020; Supriyati et al., 2021). Teachers can later use this to improve the quality of their students (Zamri & Nordin, 2015). Rasch Model analysis provides information about test-takers consistency in answering questions and the likelihood of being careless in answering, guessing, or cheating (Isnani et al., 2019; Susongko, 2016). Based on more in-depth information about the quality of their students, teachers can be helped in evaluating the learning process in the classroom (Mursidi & Soeharto, 2016; Rahmani, 2018; Suranata et al., 2018).

Based on more in-depth information about the quality of their students, educators can be helped in evaluating the learning process in the classroom (Mursidi & Soeharto, 2016; Rahmani, 2018; Suranata et al., 2018). This informs educators in improving the quality of their students (Zamri & Nordin, 2015).

The results of the Rasch Model analysis were also carried out to classify the level of mastery of each student's concepts. The results of the analysis showed that students were divided into three categories. The three categories are students with high abilities, students with medium abilities, and students with low abilities. The classification data is then used as reference to provide learning design recommendations for each category of students. These recommendations aim to assist teachers in developing future lesson plans and realizing differentiated instruction in the classroom.

Focus Group Discussion (FGD) consisting of several teachers was presented to discuss the determination of the best learning design for each category of students. This FGD activity presented teachers from senior high schools (public), teachers from faith-based high schools, and teachers from vocational high schools. The results of determining the learning design of this FGD activity can be seen in Table 2.

Table 2. Learning design recommendation.

Number	Student's Category	Learning Design Recommendation
1	Low ability	• Need a longer and more focused learning process. • Need an additional learning between students and teachers and between students and parents. • Need additional teaching materials, such as handouts with light complexity. • Must be actively invited to ask questions and discuss intensely in class.
2	Medium ability	• Learn as usual schedule. • Can be given additional learning between students and teachers. • May be given supplemental teaching materials, such as learning modules that have medium complexity. • Must be actively invited to solve problems with high intensity.
3	High ability	• Only require a short learning process. • Must be given an enrichment program (independent assignment) • May be given additional teaching materials, such as learning modules with high complexity. • Must be actively invited to solve problems with high intensity and share them with other students.

The results of learning design recommendations from the FGD are then used as references in diagnostic test instruments. This recommendation will become a teacher's reference after students have done the questions in the diagnostic test instrument developed in this study.

This research then conducts the implementation stage. The trial was performed again on 261 students who had previously collected data on migration at the school. All students are asked to take tests on diagnostic test instruments. The results of student work are then analyzed again using the Rasch Model to classify low-ability, medium-, and high-ability students.

After obtaining student ability classification data, their teachers apply learning design recommendations for each student. After implementing the learning design recommendations, all students carry out formative tests to determine the achievement of each student's learning goals. The formative test results of all students are then processed, and some information can be obtained, which can be seen in Table 3.

Table 3. Trial result data.

Students' Category	Average Score	Maximum Score	Minimum Score	Percentage of Learning Achievement
Local Students	82.20	100.00	63.00	70.70% (139 students)
Migration Students	82.25	100.00	64.00	71.43% (45 students)

Based on the trial data in Table 3, formative tests between local and migration students do not have significant differences. Through the implementation of the recommended learning design by the diagnostic test instrument, the final results between local students and migratory students no longer have significant differences. When compared to the previous one, which had a difference in the percentage of learning achievement of more than 6%, the difference in the percentage of learning achievement dropped to below 1% after implementing the diagnostic test.

The results of this trial show that the developed diagnostic test instrument can realize differentiated instruction in the classroom. Implementing differentiated instruction in the classroom gives local and migration student groups equality in learning, especially in knowledge.

The results of this study show that there is a relationship between the phenomenon of migration and differentiated instruction. In practice, differentiated instruction has succeeded in solving problems in the world of education related to the phenomenon of migration. The diagnostic test instrument developed in this study has succeeded in becoming a "bridge" between the phenomenon of migration and differentiated instruction.

The results of this study also explain that the use of diagnostic test instruments in the future will become necessary before the learning process begins. Thus, teachers have obtained a classification of each student to provide specific learning designs for each student category. As a result, differentiated instruction can be realized and can bridge the gap in fundamental knowledge in the classroom due to migration.

Conclusion

Based on the results and discussions, this study has succeeded in developing diagnostic test instruments to realize differentiated instruction that can bridge the gap in fundamental knowledge in the classroom due to migration. The differentiated instruction in the school has succeeded in creating equality in the learning process, especially in the realm of knowledge.

Nevertheless, this research still needs to be developed again. Fundamental knowledge gaps between local and migratory students can occur in other materials. Therefore, further development of diagnostic test instruments for other materials is necessary. In addition, further development can also be carried out related to different formative test instruments for each category of students' basic abilities.

References

Al Zoubi, F., Mayo, N., Rochette, A., & Thomas, A. (2018). Applying modern measurement approaches to constructs relevant to evidence-based practice among Canadian physical and occupational therapists. Implementation Science, 13(1), 1–21. https://doi.org/10.1186/s13012-018-0844-4

Alstete, J. W., Meyer, J. P., & Beutell, N. J. (2020). Enriching management learning with differentiated instruction. International Journal of Educational Management, 35(3), 640–654. https://doi.org/10.1108/IJEM-06-2020-0312

Alwi, I. (2017). Sensitivity of Mantel Haenszel Model and Rasch Model as Viewed from Sample Size. Jurnal Evaluasi Pendidikan, 2(1), 18–29. https://doi.org/10.21009/jep.021.02

Barus, G., Sinaga, J. D., & Moron, Y. D. (2019). Testing the quality of film-based creative and innovative character education test with rasch model. Jurnal Konseling Dan Pendidikan, 7(2), 71–80. https://doi.org/10.29210/136000

Bodzin, A., Hammond, T., Fu, Q., & Farina, W. (2020). Development of Instruments to Assess Students' Spatial Learning Attitudes (SLA) and Interest in Science, Technology and Geospatial Technology (STEM-GEO). International Journal of Educational Methodology, 6(1), 67–81. https://doi.org/10.12973/ijem.6.1.67

Chang, J., Li, Y., Song, H., Yong, L., Luo, L., Zhang, Z., & Song, N. (2020). Assessment of Validity of Children's Movement Skill Quotient (CMSQ) Based on the Physical Education Classroom Environment. BioMed Research International, 2020(1), 1–11. https://doi.org/10.1155/2020/8938763

Ginja, T. G., & Chen, X. (2020). Teacher Educators' Perspectives and Experiences towards Differentiated Instruction. International Journal of Instruction, 13(4), 781–798. https://doi.org/10.29333/iji.2020.13448a

Hamdu, G., Fuadi, F. N., Yulianto, A., & Akhirani, Y. S. (2020). Items Quality Analysis Using Rasch Model To Measure Elementary School Students' Critical Thinking Skill On Stem Learning. JPI (Jurnal Pendidikan Indonesia), 9(1), 61–74. https://doi.org/10.23887/jpi-undiksha.v9i1.20884

Isnani, I., Utami, W. B., Susongko, P., & Lestiani, H. T. (2019). Estimation of college students' ability on real analysis course using Rasch model. Research and Evaluation in Education, 5(2), 95–102. https://doi.org/10.21831/reid.v5i2.20924

Kamarulzaman, M. H., Azman, H., & Zahidi, A. M. (2021). The Practice of Online Differentiated Instruction and Its Impact on Motivation and Academic Performance in the Wake of Covid-19. Preprints.Org, 1(1), 1–21.

Koçak, D. (2020). Investigation of rater tendencies and reliability in different assessment methods with many facet rasch model. International Electronic Journal of Elementary Education, 12(4), 349–358. https://doi.org/10.26822/iejee.2020459464

Kopeyev, Z., Mubarakov, A., Kultan, J., Aimicheva, G., & Tuyakov, Y. (2020). Using a personalized learning style and google classroom technology to bridge the knowledge gap on computer science. International Journal of Emerging Technologies in Learning, 15(2), 218–229. https://doi.org/10.3991/ijet.v15i02.11602

Lo, C., Liang, W., Hang, L., Wu, T., Chang, Y., & Chang, C. (2015). A Psychometric Assessment of the St . George's Respiratory Questionnaire in Patients with COPD using Rasch Model Analysis. Health and Quality of Life Outcomes, 13(131), 1–16. https://doi.org/10.1186/s12955-015-0320-7

Magableh, I. S. I., & Abdullah, A. (2020). Effectiveness of differentiated instruction on primary school students' English reading comprehension achievement. International Journal of Learning, Teaching and Educational Research, 19(3), 20–35. https://doi.org/10.26803/ijlter.19.3.2

Maulana, R., Smale-Jacobse, A., Helms-Lorenz, M., Chun, S., & Lee, O. (2020). Measuring differentiated instruction in The Netherlands and South Korea: factor structure equivalence, correlates, and complexity level. European Journal of Psychology of Education, 35(4), 881–909. https://doi.org/10.1007/s10212-019-00446-4

Maulana, S., Supriyati, Y., & Astra, I. M. (2020). Development of Physics Problems for The Final Assessment in High School Levell with Rasch Model Analysis. Unnes Science Education Journal, 9(1), 22–29.

Md Yunos, J., Ibrahim Mukhtar, M., Alias, M., Lee, M. F., Tee, T. K., Rubani, S. N. K., Hamid, H., Yunus, F. A. N., Sulaiman, J., & Sumarwati, S. (2017). Validity of vocational pedagogy constructs using the rasch measurement model. Journal of Technical Education and Training, 9(2), 35–45.

Mohamad, M. M., Lisa, N., Sern, L. C., & Mohd, K. (2015). Measuring the Validity and Reliability of Research Instruments. Procedia - Social and Behavioral Sciences, 204(1), 164–171. https://doi.org/10.1016/j.sbspro.2015.08.129

Mokshein, S. E., Ishak, H., & Ahmad, H. (2019). The use of rasch measurement model in English testing. Cakrawala Pendidikan, 38(1), 16–32. https://doi.org/10.21831/cp.v38i1.22750

Mursidi, A., & Soeharto. (2016). An Introduction: Evaluation of Quality Assurance for Higher Educational Institutions using Rasch Model. Journal of Education, Teaching and Learning, 1(1), 1–6.

Nielsen, T., & Dammeyer, J. (2019). Measuring higher education students' perceived stress: An IRT-based construct validity study of the PSS-10. Studies in Educational Evaluation, 63(1), 17–25. https://doi.org/10.1016/j.stueduc.2019.06.007

Nienkerke, I. M., Thorat, A., & Patt, A. (2023). From distress migration to selective migration: Transformative effects of agricultural development on seasonal migration. World Development Perspectives, 29(1), 1–11. https://doi.org/10.1016/j.wdp.2022.100483

Onyishi, C. N., & Sefotho, M. M. (2020). Teachers' perspectives on the use of differentiated instruction in inclusive classrooms: Implication for teacher education. International Journal of Higher Education, 9(6), 136–150. https://doi.org/10.5430/ijhe.v9n6p136

Palieraki, S., & Koutrouba, K. (2021). Differentiated instruction in information and communications technology teaching and effective learning in primary education. European Journal of Educational Research, 10(3), 1487–1504. https://doi.org/10.12973/EU-JER.10.3.1487

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. Physical Review Physics Education Research, 15(2), 1–14. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Prieto, L. P., Rodríguez-Triana, M. J., Martínez-Maldonado, R., Dimitriadis, Y., & Gašević, D. (2019). Orchestrating learning analytics (OrLA): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. Australasian Journal of Educational Technology, 35(4), 14–33. https://doi.org/10.14742/ajet.4314

Rahmani, B. D. (2018). Differential Item Functional Analysis on Pedagogic and Content Knowledge (PCK) Questionnaire for Indonesian Teachers using RASCH Model. Journal of Physics: Conference Series, 948(1), 1–8.

Ramli, R., & Nurahimah, M. Y. (2020). Self-Efficacy and Differentiated Instruction: A Study among Malaysian School Teachers. Universal Journal of Educational Research, 8(4), 1252–1260. https://doi.org/10.13189/ujer.2020.080416

Şentürk, C., & Sari, H. (2018). Investigation of The Contribution of Differentiated Instruction into Science Literacy. Qualitative Research in Education, 7(2), 197–237. https://doi.org/10.17583/qre.2018.3383

Shareefa, M. (2021). Using Differentiated Instruction in Multigrade Classes: A Case of a Small School. Asia Pacific Journal of Education, 41(1), 167–181. https://doi.org/10.1080/02188791.2020.1749559

Shareefa, M., Moosa, V., Zin, R. M., Abdullah, N. Z. M., & Jawawi, R. (2019). Teachers' Perceptions on Differentiated Instruction: Do Experience, Qualification and Challenges

Matter? International Journal of Learning, Teaching and Educational Research, 18(8), 214–226. https://doi.org/10.26803/ijlter.18.8.13

Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated Instruction in Secondary Education: A Systematic Review of Research Evidence. Frontiers in Psychology, 10(1), 1–23. https://doi.org/10.3389/fpsyg.2019.02366

Smith, M. D., & Wesselbaum, D. (2020). COVID-19, Food insecurity, and migration. Journal of Nutrition, 150(11), 2855–2858. https://doi.org/10.1093/jn/nxaa270

Supriyati, Y., Susanti, D., & Maulana, S. (2021). Computer-Based Application for High School Physics Exams using IRT Model 1P. AIP Conference Proceedings, 2320(1), 1–9. https://doi.org/10.1063/5.0037567

Suranata, K., Rangka, I. B., Ifdil, I., Ardi, A., Susiani, K., Prasetyaningtyas, W. E., Daharnis, D., Alizamar, A., Erlinda, L., & Rahim, R. (2018). Diagnosis of students zone proximal development on math design instruction : A Rasch analysis. Journal of Physics: Conference Series, 1114(1), 1–7.

Susongko, P. (2016). Validation of Science Achievement Test with the Rasch Model. Jurnal Pendidikan IPA Indonesia, 5(2), 268–277. https://doi.org/10.15294/jpii.v5i2.7690

Urbański, M. (2022). Comparing Push and Pull Factors Affecting Migration. Economies, 10(1), 1–15. https://doi.org/10.3390/economies10010021

van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2019). Capturing the Complexity of Differentiated Instruction. School Effectiveness and School Improvement, 30(1), 51–67. https://doi.org/10.1080/09243453.2018.1539013

Veas, A., Castejón, J. L., Gilar, R., & Miñano, P. (2016). Validation of the Spanish adaptation of the School Atitude Assessment Survey-Revised using multidimensional Rasch analysis. Anales de Psicología, 33(1), 74–81. https://doi.org/10.6018/analesps.33.1.235271

Vinke, K., Bergmann, J., Blocher, J., Upadhyay, H., & Hoffmann, R. (2020). Migration as Adaptation? Migration Studies, 8(4), 626–634. https://doi.org/10.1093/migration/mnaa029

Whitley, J., Duquette, C., Gooderham, S., Elliott, C., Orders, S., & Klan, A. (2021). Implementation of a Differentiated Instruction Initiative: Perspectives of Leaders. Canadian Journal of Educational Administration and Policy, 196(1), 49–64. https://doi.org/10.7202/1078517AR

Xu, R. H., Wong, E. L. yi, Lu, S. Y. jun, Zhou, L. M., Chang, J. H., & Wang, D. (2020). Validation of the Toronto Empathy Questionnaire (TEQ) Among Medical Students in China: Analyses Using Three Psychometric Methods. Frontiers in Psychology, 11(1), 1–11. https://doi.org/10.3389/fpsyg.2020.00810

Yasin, R. M., Yunus, F. A. N., Rus, R. C., Ahmad, A., & Rahim, M. B. (2015). Validity and Reliability Learning Transfer Item Using Rasch Measurement Model. Procedia - Social and Behavioral Sciences, 204(1), 212–217. https://doi.org/10.1016/j.sbspro.2015.08.143

Yasin, S. N. T. M., Yunus, M. F. M., & Ismail, I. (2018). The use of rasch measurement model for the validity and reliability. Journal of Counseling and Educational Technology, 1(2), 22–27. https://doi.org/10.32698/0111

Zafiri, M., Konstantinidou, A., & Pliogou, V. (2019). The Application of Differentiated Instruction in Reading and Writing to a Boy with Autism in Early Childhood Education: A Case Study. Universal Journal of Educational Research, 7(12), 2609–2626. https://doi.org/10.13189/ujer.2019.071208

Zamri, A., & Nordin. (2015). Modeling a Multiple Choice Mathematics Test with the Rasch Model. Indian Journal of Science and Technology, 8(12), 1–6. https://doi.org/10.17485/ijst/2015/v8i12/70650