RedWhitepress https://series.gci.or.id

Global Conferences Series:
Social Sciences, Education and Humanities (GCSSSEH), Volume 8, 2021
International Conference on Special Education In South East Asia Region 11th Series 2021
DOI: https://doi.org/10.32698/GCS-11512

Item response theory approach in the preparation of instruments for identification of mathematics learning difficulties

Indra Praja Kusumah[a], Budi Susetyo[a], Iding Tarsidi[a], Yubali Ani[b]
[a]Universitas Pendidikan Indonesia, Bandung, Indonesia,[b]Universitas Pelita Harapan, Tanggerang , Indonesia
E-mail: indra.kusumah@upi.edu

Abstract: The difficulty of learning mathematics can be experienced by anyone, from elementary school students to students in college. Difficulty learning mathematics in students if not detected early will have a special detrimental impact on Primary teacher students who will later become teachers in elementary school. Basic mathematics is a basic competency that must be mastered for Primary teacher students. In fact, in the field, there are still many Primary teacher students who have difficulty learning mathematics, especially basic mathematics. However, research on the instrument identification test of the difficulty of learning mathematics in students in basic mathematics courses is still very little especially with the approach of item response theory. Some studies still use the classic theory that is sample bound. The purpose of this study is to (1) describing the multiple-choice test instruments with the three-parameter model item response theory approach and (2) analyzing whether the instrument test multiple choice with the IRT 3PL theory approach can identify the difficulty of learning the mathematics of Primary teacher students in basic mathematics courses. The research method used is quantitative research and *development* (R&D) with a total of 250 Primary teacher students at a private university in Tangerang. The results showed multiple-choice test instruments with the IRT 3PL approach can identify students who have difficulty learning mathematics in basic mathematics courses.

Keywords: Item response theory, mathematics learning difficulties, instruments identification

INTRODUCTION

Learning difficulties are non-permanent learning disorders and are usually influenced by external factors such as emotional state, education, or the environment. Learning difficulties do not correlate or indicate a certain level of intelligence but indicate that a person faces certain difficulties in learning[1] Learning difficulties are different from learning disabilities. Learning disabilities under the Individual Disability Education Act are defined as disorders in one or more basic psychological processes involved in the ability to understand or use language (oral or written), in which disorders manifest in the inability to listen, think, speak, read, write, spell or perform mathematical calculations [2] This disorder is located in the neurological system resulting in disruption of the development, integration, and/or demonstration of verbal and/or nonverbal abilities and is permanent [1] The difficulty of learning mathematics in this study is not dyscalculia because (1) there is still much debate among experts as to whether dyscalculia belongs to learning disability or learning difficulty that is not handicap-related [3], [4]; (2) refers to DSM-V, dyscalculia is a disorder that causes a person to have difficulty learning concepts related to numbers or use symbols and functions to perform mathematical calculations [4], [5] The difficulty of learning mathematics is defined as the condition in which an individual fails to obtain adequate proficiency in basic mathematical ability or has lower academic achievement than expected[6]–[9]. Similarly, Lithner [3] defines the difficulty of learning mathematics as a condition where students have difficulty in understanding content and difficulties in mathematical thinking (problem-solving, proof and proving, reasoning, and modeling). The difficulty of learning mathematics is not only experienced by elementary or secondary school students but can also be experienced by students in college [3], [10], [11]. Difficulty learning mathematics if not recognized from the beginning will have a special detrimental impact on Primary teacher students who will later become teachers in elementary schools. Basic mathematics is a mandatory basic competency that must be mastered for all Primary teacher students. There are still many Primary teacher students who have difficulty learning mathematics, especially basic mathematics [12], [13] But unfortunately, the research of identification test instruments difficulty learning mathematics in students in basic mathematics courses is still very little especially with the approach of item response theory. Some studies still use classical exam theory that has a deficiency that is the item difficulty of the question item still depends on the level of ability of the subject or sample bound [14] As a result, the measuring instrument can not measure the ability of the respondent actually because the measuring instrument is tested to a respondent with high capability as if it has a low level of difficulty and vice versa.

Based on this, this study aims to (1) describe the number of multiple-choice test instruments with the three-parameter model item response theory approach and (2) analyze whether the instrument of the multiple-choice test with the IRT L3P theory approach can identify the difficulty of learning mathematics primary teacher students in basic mathematics courses.

METHOD

The research method used is quantitative R&D. Research and development (R&D) according to Sugiyono [15, p. 30] is "a scientific way to research, design, produce and test the validity of the resulting product". Silalahi [16] added R&D methods focusing not on formulating or testing theories but on developing effective products. In education, the products in question can be textbooks, films for learning, computer software, test/evaluation instruments, teaching methods, and educational programs Borg and Gall, 1998, dalam [15, p. 28] This study was conducted with respondents of 250 Primary teacher students of the faculty of education at a private university in Tangerang Banten.

RESULT AND DISCUSSION

Mathematics Learning Difficulties

Setianingrum, R., Syamsuri, S., &Setiani, Y. [17] defines the difficulty of learning mathematics as "the student's inability to solve a math problem". More specifically, the difficulty of learning mathematics is defined as a condition in which an individual fails to obtain sufficient proficiency in basic mathematical ability or has lower academic achievement than expected but the cause is not due to intellectual disability or other internal factors [6]–[9]. Similarly, Lithner, J. [3] defines the difficulty of learning mathematics as "a condition where students have difficulty in understanding content as well as difficulties in mathematical thinking (problem-solving, proof and proving, reasoning, and modeling)that may be due to previous learning experiences".

Lithner's opinion is reinforced by Jamaris, M. [18, p. 129] who argues that "the difficulty of learning mathematics is a difficulty in thinking and processing mathematically (deductive, rational, and logical thinking) resulting from external factors characterized by low learning outcomes or far below the individual's potential. Scherer, P., Beswick, K., DeBlois, L., Healy, L., &Opitz, E. [19] argues that children with difficulty learning mathematics are "children who have poor achievements in mathematics subjects resulting from the failure of children in mastering basic mathematical concepts and abilities that are not connected due to certain disabilities and have cognitive abilities within limits or above normal".

So in general, the term difficulty in learning mathematics refers to children whose academic achievements are poor or below average due to various factors such as poor teaching or environmental factors. This difficulty in learning mathematics is not caused or related to the disability or intelligence skills concerned. Individuals with difficulty learning mathematics sometimes perform well in other fields but they are very difficult in mathematics.

Basic College Mathematics

Amir, M., &Prasojo, B., [20, pp. 1–3] in his book "Basic Mathematics Textbook" writes eight materials taught in basic mathematics courses namely real number systems, sets, linear equations and inequalities, functions, matrices, limits and continuity, derivatives and integrals. A team of lecturers in basic mathematics courses at the University of Semarang [21] described basic mathematics courses discussing real number systems, inequality, absolute values, square and square roots, cartesian and polar coordinates, graphs, linear equation systems, functions and limits, derivatives, derivative applications, integrals, and integral applications. Yahya, Y., et al., [22, pp. vi–viii] mentions the material in basic mathematics in higher education consisting of "sets, numbers, vectors and matrices, functions, limits, and continuity of functions, derivatives, and integrals".

So based on the explanation above basic mathematics in college is an extension of the math materials that have been obtained in high school. The subjects to be used in this study based on the above explanations are set, real number system, linear equations and inequalities, absolute values, functions and graphs, limits, derivatives and applications, and their integrals and applications.

Item Response Theory Three Parameters Logistic (IRT 3PL)

The Response Theory (IRT) item is basically used to address the weaknesses of classical exam theory. IRT analyzes the item difficulty of items directly related to the characteristics of the item. This is in contrast to the classic test theory that analyzes the difficulty of items directly related to the subject's abilities. In its mathematical model, IRT illustrates that the probability of the subject answering the item correctly largely depends on the ability of the subject and the characteristics of the item[23, p. 1] IRT basically has a rule where the subject's chances of answering correctly, item parameters, and ability parameters are associated with a mathematical formula model that must be adhered to either by the test taker group or the test item group [24] The rules in this IRT cause the difficulty of the item to be independent of the item difficulty of the item is not affected by the level of ability of the subject meaning that the same problem item has the same level of difficulty even though it is answered by subjects of different levels of ability[14, p. 61]

IRT has item characteristics so it is necessary to determine the model of item characteristics used in analyzing the test instrument to be compiled. Based on the logistics model IRT can be in the form of (1) Logistics one parameter (L1P) where the parameter of the item is only the level of difficulty, (2) Logistics two parameters (L2P) where the parameter of the item level of difficulty and power is different, (3) Logistics three parameters (L3P) where the parameter items the level of difficulty, power difference, and guess[14], [24]. This study only used three parameters logistics model. IRT in its use has four requirements that must be met, namely (1) logistics model, (2) unidimensional, (3) group invariance, and (4) local independence [14, p. 64]

1. Logistics Model

   The purpose of the model check will be used to see the match of empirical data on each test item with the logistics model. Check the model by comparing values $\chi^2\ (p - value) > 0.05$ [25]. There are three logistics models in the IRT approach, namely (1) One-parameter logistics model in the form of item difficulty, (2) Two-parameter logistics model in the form of difficulty and power level, and (3) three-parameter logistics model in the form of difficulty, power difference, and guessing. Model logistics used in this research are three-parameter logistics (3PL). DeMars [25] says "among the three-parameter model dichotomy models are the most suitable models for multiple choice-shaped items".
2. Unidimensionality

   Unidimensional tests are conducted to ensure variables to be measured with the test device are in the same dimension. In the study, the unidimensional test to ensure that the measuring instrument given to respondents was only to measure basic mathematical skills rather than others. Technical analysis used using factor analysis is analyzing the value of *eigenvalue* factor with the provision of *eigenvalue* value of the first factor should be greater than *eigenvalue* the second factor and *eigenvalue* the second factor the value is almost the same as the next factor [14, p. 65] Susetyo in detail [14, p. 69] describes the steps of testing unidimensional requirements i.e. "calculation of KMO and Bartlett's Tests, *anti-image* correlation and *eigenvalue* factor. Testing can be done using SPSS computer programs".
3. Invariance

   The group's invariance requirements to ensure that the measuring instruments compiled have the same item characteristics when given to all subgroups of respondents [14, p. 65] In this study with a three-parameter model (item difficulty, item discrimination, and pseudo guessing) then the examination was conducted against the item parameters (item difficulty, item discrimination, and pseudo guessing) and the ability parameters of the respondent. Analysis of invariance tests on item parameters is conducted by correlation test between high group items and low group items on each parameter i.e., invariance test at item difficulty, item discrimination, and past guesses by using scatter plot to see the spread of its value. While the analysis of invariance test on the parameters of the ability of respondent through the correlation between the ability of high groups and low groups. After that, check by using a scatter plot to find out the spread of its value [14, p. 72].

4. Local Independence

Local independence testing to ensure the probability of answering correctly for different items at the same location with the same capability is independent of one item against another test item [14, p. 66] According to Sudaryono [24] local independence is divided into two, namely "local independence to the response of respondent and local independence to test points". Local independence to the response means that the participant's fault is not influenced by the correctness of the other respondent in answering the same item. While local independence of the item means that one of the respondents answered an item is not affected by the correctness of the test taker in answering the other item [24] Analysis of local independence tests can be done in two ways, namely by excretion through probability formulas and statistically with chi-squared dependency test.

Result

The instruments used in this study are 50 items of multiple-choice questions with five answer options. The material used is set, real number system, linear equations and inequalities, absolute values, functions and graphs of functions, linear line equations, limits and continuity, derivatives and its' applications, and integrals and its' applications. The stages performed by researchers in this study are content validity test, reliability test, assumption test, and final analysis.

1. Content Validity and Reliability Test

The content validity test method used by researchers is to see the match of test items with indicators. The measuring instrument is said to be valid with this method if the percentage of matches with the indicator is greater than 50% [14, p. 116] Validity results show 50 questions declared escaped by four validators

The reliability test method used by researchers is an internal consistency reliability test with the *Kuder-Richardson* technique. (KR_{20}) A device is said to be reliable when the correlation coefficient value *is Kuder-Richardson* or $\rho_{KR_{20}} > 0.70$[14, p. 152] Based on figure 1, the measuring instrument is said to be reliable due to the coefficient value.$0.928 > 0.70$

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.928	.927	50

Figure 1. Reliability Test Result

2. Test Assumptions

2.1 3PL Model Fit

L3P model match test results found eight items with values$\chi^2 \leq 0.05$ [26] namely point 6, 9,10,33, 34, 37, 40,44 (see table 1). These eight items will be removed from the next analysis process.

Table 1 L3P Model Matches

No.	χ^2	No.	χ^2	No.	χ^2	No.	χ^2	No.	χ^2
no1	0.303	no11	0.664	no21	0.548	no31	0.161	no41	0.715
no2	0.125	no12	0.630	no22	0.147	no32	0.291	no42	0.528
no3	0.746	no13	0.492	no23	0.897	no33	0.028	no43	0.145
no4	0.063	no14	0.865	no24	0.411	no34	0.017	no44	0.011
no5	0.125	no15	0.485	no25	0.435	no35	0.452	no45	0.732
no6	0.006	no16	0.051	no26	0.489	no36	0.189	no46	0.161
no7	0.996	no17	0.381	no27	0.860	no37	0.011	no47	0.794
no8	0.139	no18	0.072	no28	0.803	no38	0.972	no48	0.616
no9	0.017	no19	0.343	no29	0.807	no39	0.735	no49	0.888
no10	0.020	no20	0.557	no30	0.295	no40	0.011	no50	0.922

2.2 Unidimensionality

The result of the unidimensionality test shows the value of eigenvalue the first factor > the value of eigenvalue the second factor, as well as the value of eigenvalue the second factor > the value of eigenvalue the third factor, and so on. This indicates that there is only one dominant factor in this test instrument, namely the first factor [14, p. 65] So it can be concluded that this test instrument passed the unidimensional test

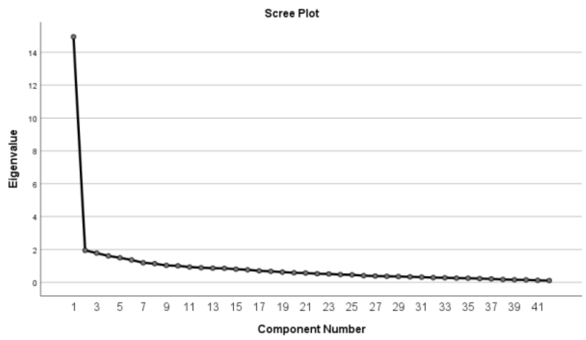

Figure 2 Scree plot Eigenvalue

3. Invariance

Invariant testing is performed against the parameters of ability and parameters of items (item discrimination, item difficulty, and pseudo guessing). The first stage of the test of invariance of capability parameters (theta) by looking at the correlation between the value theta respondents of the low group with the value of theta respondents of the high group. The calculation resulted in a correlation value of 0.844 so that the correlation between the ability of respondents of the high group with the ability of respondents of the low group is very high (figure 3).

The second stage of the test of invariance of item parameters is by looking at the correlation between the high group and the low group on item discrimination, item difficulty, and pseudo guess. The test results resulted in a correlation value of 0.747 (figure 4) between the item discrimination groups of high and the item discrimination of the low group. So it can be said that the item discrimination between the high and low groups has a fairly high correlation. Testing correlation item difficulty between high and low groups is quite high, this can be seen from the correlation value of 0.742 (figure5). The result of correlation testing pseudo guessing between high and low groups gives the same results. The correlation value is 0.753 or indicates a correlation between the two groups. The conclusion of the basic mathematical test device in this study meets the requirements of the invariance test meaning that the basic mathematical test device has the same item characteristics (invariant) for all subgroups or homogeneous if all respondents in that subgroup have the same ability.

Correlations Parameter Kemampuan

		theta_kel. rendah	theta_kel. tinggi
theta_kel.rendah	Pearson Correlation	1	.844**
	Sig. (2-tailed)		.000
	Sum of Squares and Cross-products	1.929	2.193
	Covariance	.029	.033
	N	68	68
theta_kel.tinggi	Pearson Correlation	.844**	1
	Sig. (2-tailed)	.000	
	Sum of Squares and Cross-products	2.193	3.503
	Covariance	.033	.052
	N	68	68

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 3 Correlation Parameter Capabilities

Correlations Parameter Butir Daya Beda

		dayabeda_kel.rendah	dayabeda_kel.tinggi
dayabeda_kel.rendah	Pearson Correlation	1	.747**
	Sig. (2-tailed)		.000
	Sum of Squares and Cross-products	1.260	1.913
	Covariance	.063	.096
	N	21	21
dayabeda_kel.tinggi	Pearson Correlation	.747**	1
	Sig. (2-tailed)	.000	
	Sum of Squares and Cross-products	1.913	5.211
	Covariance	.096	.261
	N	21	21

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4 Correlations Item Parameter Item Discrimination

Correlations Parameter Butir Tingkat Kesukaran

		sukar_kel.rendah	sukar_kel.tinggi
sukar_kel.rendah	Pearson Correlation	1	.742*
	Sig. (2-tailed)		.045
	Sum of Squares and Cross-products	17.454	21.007
	Covariance	.873	1.050
	N	21	21
sukar_kel.tinggi	Pearson Correlation	.742*	1
	Sig. (2-tailed)	.045	
	Sum of Squares and Cross-products	21.007	129.461
	Covariance	1.050	6.473
	N	21	21

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 5 Correlations Item Parameter Difficulty Level

Correlations Parameter Butir Tebakan

		tebakan_kel.rendah	tebakan_kel.tinggi
tebakan_kel.rendah	Pearson Correlation	1	.753**
	Sig. (2-tailed)		.000
	Sum of Squares and Cross-products	.032	.055
	Covariance	.002	.003
	N	21	21
tebakan_kel.tinggi	Pearson Correlation	.753**	1
	Sig. (2-tailed)	.000	
	Sum of Squares and Cross-products	.055	.166
	Covariance	.003	.008
	N	21	21

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 6 Correlations Item Parameter Pseudo Guess

3.1 Local Independence

Local independence test is performed by looking at the covariance variance values of ten interval groups (100 first responders) capability value (theta) after being sorted from highest to lowest values indicating covariance values close to zero [14, p. 76] The test result is seen in figure 7, it appears that the covariance value is decreasing and close to zero. So, it can be concluded on this test device on the same ability (theta) then the probability of answering correctly $P(\theta)$ for different items is independent of one test item against another test item.

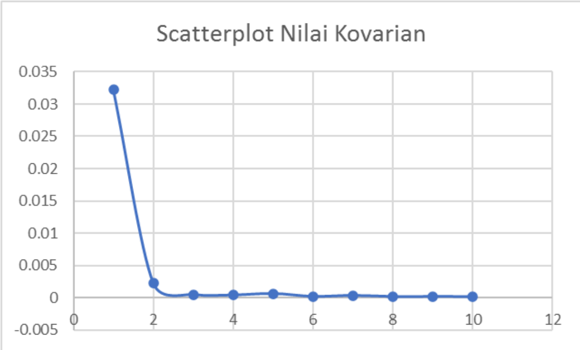

Figure 7 Scatterplot of Covariance Values

In the conclusion, a basic mathematical test device with multiple-choice forms with five answer options meets the assumption requirements of IRT 3PL.

4. Identification of Mathematics Learning Difficulties based on thetas' value.

The principle of the preparation of cognitive test devices with IRT approach will result in sample *bound-free* test points and produce capability parameter values (theta) that are close to the actual respondent's ability value [14], [25], [27]–[31] This principle is also used by some researchers in the preparation of cognitive test devices to uncover the abilities of respondents such as Amelia & Kriswantoro [32]Lonnemann & Hasselhorn [33] Zhao [34] Zhao [34] used the average respondent's theta score as a benchmark to analyze math skills and math learning difficulties in students in his research, while Lonnemann &Hasselhorn [33] used a low group (27% of the total respondents) to determine students who had low math skills in his research. Thus the method of determining the difficulty of learning mathematics primary teacher students in basic mathematics courses in the study combined Lonnemann & Hasselhorn [33] and Zhao [34] Based on this, 34 Primary teacher students have difficulty learning mathematics in basic mathematics courses

CONCLUSION

The preparation of mathematical ability test instruments with multiple choice with five answer options with the IRT 3PL approach must meet the assumption of IRT 3PL i.e. unidimensionality, invariance (item parameters and capability parameters) before further analysis is launched. The mathematical ability test instruments in this study have already met the requirements of the assumption test and can be used to identify students who have difficulty learning mathematics by looking at their theta scores.

REFERENCES

M. J. Aftab, T. M. Khan, and T. Hussain, "Validation of Screening Checklist for Learning Difficulties in Mathematics," *Bull. Educ. Res.*, vol. 39, no. 2, pp. 119–126, 2017.

R. Reid and T. O. Lienemann, *Strategy Instruction for students with learning disabilities*. New York: The Guilford Press, 2006.

J. Lithner, "University Mathematics Students' Learning Difficulties," *Educ. Inq.*, vol. 2, no. 2, pp. 289–303, Jun. 2011, doi: 10.3402/edui.v2i2.21981.

N. Soares, T. Evans, and D. R. Patel, "Specific learning disability in mathematics : a comprehensive review," *Transl. Pediatr.*, vol. 7, no. 1, pp. 48–62, Jan. 2018, doi: 10.21037/tp.2017.08.03.

D. Penesetti, "What Is Specific Learning Disorder?," *American Psychiatric Association*, 2018. https://www.psychiatry.org/patients-families/specific-learning-disorder/what-is-specific-learning-disorder (accessed Nov. 13, 2020).

E. M. Yeni, "Kesulitan Belajar Matematika Di Sekolah Dasar," *JUPENDAS*, vol. 2, no. 2, pp. 1–10, 2015.

E. Rosita, "Diagnosis kesulitan belajar," *Univ. Negeri Yogyakarta*, vol. I, no. 09403241034, pp. 19–26, 2010.

P. Westwood, *What teachers need to know about: Learning difficulties*. Victoria: ACER Press, 2008.

P. Westwood, *Commonsense Methods for Children with Special Needs*, Sixth. New York: Routledge, 2015.

S. T. Kane, S. Roy, and S. Medina, "Identifying College Students At-Risk for Learning Disabilities:

Evidence for Use of the Learning Difficulties Assessment in Postsecondary Settings," *J. Postsecond. Educ. Disabil.*, vol. 26, no. 1, pp. 21–33, 2013.
S. Klymchuk, T. Zverkova, N. Gruenwald, and G. Sauerbier, "University Students' Difficulties in Solving Application Problems in Calculus," *Math. Educ. Res. J.*, vol. 22, no. 1, pp. 81–91, 2010, doi: 10.1007/BF03217567.
A. N. K. Rosyidah, M. A. Maulyda, and I. Oktaviyanti, "Miskonsepsi Matematika Mahasiswa PGSD Pada Penyelesaian Operasi Hitung Bilangan Bulat," *J. Ilm. Kontekst.*, vol. 2, no. 01, pp. 15–21, 2020, doi: 10.46772/kontekstual.v2i01.244.
J. P. Purwaningrum and H. S. Bintoro, "Miskonsepsi matematika materi bilangan pada mahasiswa calon guru sekolah dasar," in *Prosiding Seminar Nasional MIPA 2018*, 2018, no. November, pp. 173–180.
B. Susetyo, *Prosedur Penyusunan dan Analisis Tes untuk Penilaian Hasil Belajar Bidang Kognitif*. PT Refika Aditama, 2015.
Sugiyono, *Metode Penelitian dan Pengembangan (Research and Development/ R&D)*. Bandung: Alfabeta, 2019.
A. Silalahi, "Development Research (Penelitian Pengembangan) dan Research & Development (Penelitian & Pengembangan) Dalam Bidang Pendidikan/Pembelajaran," *Res. Gate*, no. July, pp. 1–13, 2018, doi: 10.13140/RG.2.2.13429.88803/1.
R. S. Setianingrum, S. Syamsuri, and Y. Setiani, "Analyzing students' learning difficulties in algebra," *J. Mat. dan Pembelajaran*, vol. 8, no. 1, pp. 19–34, Jun. 2020, doi: https://doi.org/10.24252/mapan.2020v8n1a2.
M. Jamaris, *Anak Berkebutuhan Khusus*. Jakarta: Penerbit Ghalia Indonesia, 2018.
P. Scherer, K. Beswick, L. DeBlois, L. Healy, and E. M. Opitz, "Assistance of Students with Mathematical Learning Difficulties — How Can Research Support Practice ? — A Summary," in *Proceeding of the 13th International Congress on Mathematical Education*, 2017, pp. 249–259, doi: 10.1007/978-3-319-62597-3.
M. F. Amir and B. H. Prasojo, *Matematika Dasar*. Sidoarjo: UMSIDA PRESS, 2016.
Tim Dosen, *Bahan Ajar Matematika Dasar*, Edisi Revi. Semarang: Universitas Negeri Semarang, 2012.
Y. Yahya, D. S. H. S, and A. S, *Matematika Dasar Perguruan Tinggi*, Kedua. Bogor: Tim Ghalia Indonesia, 2013.
H. Retnawati, *Teori Respons Butir dan Penerapannya*. Yogyakarta: Nuha Medika, 2014.
Sudaryono, "Implementasi Teori Responsi Butir pada THB Sekolah .pdf," *J. Pendidik. dan Kebud.*, vol. 17, no. 6, 2011.
C. DeMars, *Item Response Theory: Understanding Statistics Measurement*. New York: Oxford University Press Inc, 2010.
R. K. Hambleton and H. Swaminathan, *Item Response Theory: Principles and Applications*, vol. 9, no. 3. New York: Springer Science+Business Media, 1985.
Sudaryono, *Teori Responsi Butir*, Pertama. Yogyakarta: Graha Ilmu, 2013.
F. B. Baker and S.-H. Kim, *The Basics of Item Response Theory Using R*. USA: Springer International Publishing, 2017.
W. J. van der Linden, *Handbook of item response theory, volume three applications*. New York: CRC Press, 2016.
S. E. Embretson and S. P. Reise, *Item Response Theory for Psychologists*. London: Lawrence Erlbaum Associates, Publishers, 2000.
A. A. Bichi and R. Talib, "Item Response Theory: An Introduction to Latent Trait Models to Test and Item Development," *Int. J. Eval. Res. Educ.*, vol. 7, no. 2, p. 142, 2018, doi: 10.11591/ijere.v7i2.12900.
R. N. Amelia and Kriswantoro, "Implementasi Item Response Theory sebagai Basis Analisis Kualitas Butir Soal dan Kemampuan Kimia Siswa Kota Yogyakarta," *J. Kim. dan Pendidik. Kim.*, vol. 2, no. 1, pp. 1–12, 2017.
J. Lonnemann and M. Hasselhorn, "Assessing Mathematical Competence and Performance: Quality Characteristic, Approaches and Research Trends," in *International Handbook of Mathematical Learning Difficulties: From the Laboratory to the Classroom*, A. Fritz, V. G. Haase, and P. Räsänen, Eds. New York: Springer International Publishing, 2019, pp. 633–651.
N. Zhao, "Mathematics learning performance and Mathematics learning difficulties in China," Universiteit Gent, 2011.