# Certificate

4303/UN.48/LL/2021

Awarded to

## SYARIFA WAHIDAH AL IDRUS

## As Presenter

Title of Paper: **"GEN-Z PRE-SERVICE SCIENCE TEACHER STEM APPLICATIONS DURING ONLINE TEACHING: A RASCH ANALYSIS"**

in The International Seminar, organized by Doctor of Education Study Program, Postgraduate, Ganesha University of Education In Collaboration with Perkumpulan Program Studi Doktor Pendidikan

*"Strengthening Educational leadership, Communication, Creative Thinking, and Technology Literacy in Digital Era"*

Bali - Indonesia 27th September 2021.

Rector
of Ganesha University of Education,

Prof. Dr. Nyoman Jampel M.Pd.
NIP. 195910101986031003

Chair
Perkumpulan Program Studi Doktor Pendidikan,

Prof. Dr. Wachidi, M.Pd.
NIP. 195201111985031001

# GEN-Z PRE-SERVICE SCIENCE TEACHER STEM APPLICATIONS DURING ONLINE TEACHING: A RASCH ANALYSIS

**Syarifa Wahidah Al Idrus[1]**

Science Education, Ganesha University of Education, Singaraja, @unram.ac.id

**Abstract**

This study aims to delineate the STEM application of Gen Z pre-service science teachers in teaching science during the covid-19 pandemic. There are 140 participants taking part in the online survey. Their STEM application was assessed using a validated instrument for the Indonesian teacher context. STEM application, in this study, is divided into seven categories as the result of combining each STEM component including SAp-ST, SAp-SE, SAp-SM, SAp-STE, SAp-STM, SAp-SEM, SAp-STEM), where S, T, E, M refers to Science, Technology, Engineering, and Mathematics respectively. Based on Rasch analysis, most pre-service teachers who apply STEM teaching integration are at a moderate level. Interestingly, SAp-SEM is the most difficult combination to apply, higher than SAp-STEM application. It is also vital to note that their STEM application is not influenced by gender and specialization referring to careful analysis using the Chi-square test to each item based on the Rasch model.

Keywords: STEM education, Gen Z, Covid-19, online learning, Rasch model

## 1. Introduction

STEM education has been applied as the core agenda in many countries in the world. This is form of responses toward the rapid development of technology, for instance Industrial Revolution 4.0. Upon talking about education, children should be given strong understanding about the relationship of science learning with engineering and technology to have a meaningful learning (Elayyan, 2021; Yahya & Hashim, 2021). The importance of implementing STEM education also pivots on the characteristics of students from early childhood education until university students who are generation Z (Gen Z) and alpha generation. Their methods of perceiving knowledge tend to better different from previous generation, strengthening the crucial of integrating technology in learning (Azman et al., 2021; Manzoni, Caporarello, Cirulli, & Magni, 2021). Other than that, current situation such as covid-19 pandemic, the issue of global warming and many crucial issues, it encourages the government should go extra miles to inculcate future scientists and engineers (Zahidi et al., 2021).

Currently, Indonesia does not apply STEM education and its descendents such as STEAM (adding art component), STREAM (including religion specialization). STEM education in Indonesia is just a research topics or certain projects for researchers (Bogusevschi, Muntean, & Muntean, 2020; Khotimah, Adnan, Ahmad, & Murtiyasa, 2021). However, the application of STEM education will be near future considering the good movement of Indonesian curriculum which can transform quickly based on current trends and condition. Indonesia Ministry of Education (MoE) also tend to improve curriculum to be better through many appropriate policy.

The application of STEM education will depend on the teachers because they are the factors who will directly implement STEM curriculum. However, current teachers who are mostly millennials teachers and previous generation, it seems uneasy to apply something new. The hope can be given to the native of technology, Gen Z. They are also popular for other names such as Google generation, Viral generation, Internet generation etc (Poláková & Klímová, 2019). Their learning characteristics are unsuitability of traditional methods of teaching (Szymkowiak, Melović, Dabić, Jeganathan, & Kundi, 2021) and it potentially related to their teaching style.

Even a plethora of studies was conducted in the field, there is still lack of studies in the field of teacher development and this study is part of bigger studies in the field. When considering this study, careful analysis of wide range of literature is needed. Some meta analysis study which analyzing many previous studies such as (Mustafa, Ismail, Tasir, & Mohamad Said, 2016) evaluates a total 18 studies, (Belland, Walker, Kim, & Lefler, 2017) synthesizes 144 studies, (Iswadi, Syukri, Soewarno, Yulisman, & Nurina, 2020) reviews 16 articles and 5 chapters to explain development of STEM-TPACK scale and (Yildirim, 2016) analyze 33 studies etc.

In the current study, the analysis will depend on Rasch measurement model. It, another name of one parameter Item Response Theory (IRT), is also widely used to establish validity and reliability by reporting detailed analysis such as unidimensionality, reliability, separation, rating scale calibration, item fit statistics, Differential Item Functioning etc (William J. Boone, 2016). Rasch measurement model are based two theorems: 1) A person who is more capable has a higher probability of correctly responding to all the items provided. 2). An easier item is more likely to be answered correctly by all respondents or test- takers (Bond & Fox, 2015; Linacre, 1999). Rasch model has been applied in science education studies for instrument validation and data analysis (Jin, Rodriguez, Shah, & Rushton, 2020; Qudratuddarsi, Sathasivam, & Hutkemri, 2019; Romine, Schaffer, & Barrow, 2015; Wind & Gale, 2015).

Considering the importance of STEM education, teacher roles, research gap and Rasch model, we formulated to research questions: 1) To identify level of STEM applications of Gen-Z Pre-Service Science Teacher, 2) To identify the differences of STEM applications of Gen-Z Pre-Service Science Teacher based on gender and science specialization.

## 2. Methods

This study is quantitative study where data are in the form of number which is the result of survey research. In this study, we directly ask the participants to fill in the google form without any intervention made (Creswell, 2012; Qudratuddarsi et al., 2019). We use online google form as the effect of government rules to control movement by conducting online education (Sukendro et al., 2020). Even there is tendency to have lower response rate, this method is enhanced by direct request using social media such as WhatsApp from Facebook to gain more responses from targeted respondents (Zuidgeest, Hendriks, Koopman, Spreeuwenberg, & Rademakers, 2011).

### ***Sample of the study***

The sample are pre-service teacher who teaching during covid-19 pandemic. In Indonesian universities, prospective teachers are given a chance to directly teaching their specialization for some months depending on university and school agreement. In that phase, due to covid-19 pandemic, they also have to carry online teaching as the schools do as the responses of limiting human movements. They are 140 students with 48 male (34.28%) participants and 92 female (65.71%) participants. They are majoring science 32 participants (22.86%), chemistry 52 partcipants (37.14%), physics 34 participants (24.28%) and biology (15.71%).

Table 1. The Sample of the study

| | N | Percentage |
|---|---|---|
| Area of specialization | 140 | 100% |
| Science | 32 | 22.86% |
| Chemistry | 52 | 37.14% |
| Physics | 34 | 24.28% |
| Biology | 22 | 15.71% |
| Gender | 142 | 100 % |
| Male | 48 | 34.28 % |
| Female | 92 | 65.71 % |

***Instrument***

The instrument has been developed by previous researcher for Indonesian context (Wahono & Chang, 2019) and has been applied in some studies (Parmin, Saregar, Deta, & El Islami, 2020; Wahono & Chang, 2019). The instrument has 7 sub domain namely Science-Technology (SAp-ST), Science-Engineering (SAp-SE), Science-Mathematics (SAp-SM), Science-Technology-Mathematics (SAp-STM), Science-Technology-Engineering (SAp-STE), Science-Engineering- Mathematics (SAp-SEM), and Science-Technology-Engineering-Mathematics (SAp-STEM). In previous studies, the instrument has good validity and reliability index, indicating an appropriate measure to reveal the STEM application of Gen Z pre-service science teachers in teaching science during the covid-19 pandemic.

***Reliability and Separation***

Reliability is the degree to which an instrument consistently give a similar result among numerous administration (Fraenkel & Wallen, 2009; Shultz, Whitney, & Zickar, 2014).To measure reliability, this study applied Cronbach's alpha internal consistency to elicit the correlation between a score of an individual item in the test and the total gained score for all items (Chua, 2013). Other measures to report reliability are item and person reliability, person reliability elicits the stability of student responses in each instrument, while item reliability elicits the stability of item score (Sumintono & Widhiarso, 2015). The minimum score for each reliability (Cronbach alpha, item and person) is 0.65 (Adams, Chuah, Sumintono, & Mohamed, 2021; DeVellis, 2012), and this study found reliability in the range of 0.92-0.96, delineating an excellent score. Another result to consider is separation either item or person which should be more than 1.5 to be considered acceptable (Suryadi, Hayat, Dwirifqi, & Putra, 2021; Tennant & Conaghan, 2007). Separation of STEM applications in the context of pre-service science teachers during pandemic teaching for both item and person is 3.78 and 3.45 respectively. The result indicates the ability of instruments to distinguish item and respondents into some acceptable groups (Iseppi et al., 2021).

Table 2. Reliability and Separation

| | Value |
|---|---|
| **Reliability** | |
| Cronbach | 0.96 |
| Item | 0.92 |
| Person | 0.93 |
| **Separation** | |
| Item | 3.78 |
| Person | 3.45 |

***Item Fit Statistics***

As the proof of construct validation, mean square (MNSQ): 0.5 <MNSQ <1,5 (b) the value of tolerated infit and outfit Z- Standard (ZSTD): -2.0 <ZSTD <+2,0 (c) the value of accepted Correlation Points (Pt Mean Corr) must be positive value (W.J Boone, Staver, & Yale, 2014; Sadhu & Laksono, 2018). This analysis is very crucial as the strength of Rasch model compared to the analysis using Classical Test Theory (CTT) (Hidayat, Qudratuddarsi, Mazlan, & Zeki, 2021). This study found that each item fulfilled the criteria well, indicating the instrument can fit Rasch measurement model very well. Even there are some items violate the acceptable score, they are never in the same items all together.

Table 3. Item fit statistics of STEM application instrument

| **Item** | **MNSQ** | | **ZSTD** | | **Point Mea Corr** |
|---|---|---|---|---|---|
| SAp-ST1 | 1.21 | 1.09 | 1.7 | 0.7 | 0.61 |
| SAp-ST2 | 1.17 | 1.10 | 1.4 | 0.8 | 0.63 |
| SAp-ST3 | 1.51 | 1.58 | 3.6 | 3.0 | 0.47 |
| SAp-ST4 | 1.40 | 1.87 | 3.0 | 5.1 | 0.52 |
| SAp-SE1 | 1.19 | 1.10 | 1.5 | 0.7 | 0.57 |
| SAp-SE2 | 1.22 | 1.74 | 1.8 | 4.7 | 0.60 |
| SAp-SE3 | 1.10 | 1.15 | 0.9 | 1.1 | 0.64 |
| SAp-SM1 | 0.89 | 0.84 | -0.9 | -1.1 | 0.68 |
| SAp-SM2 | 0.86 | 0.83 | -1.2 | -1.3 | 0.72 |
| SAp-SM3 | 0.89 | 0.81 | -0.9 | -1.3 | 0.66 |
| SAp-STE1 | 0.98 | 1.35 | -0.1 | 2.5 | 0.64 |
| SAp-STE2 | 0.72 | 0.71 | -2.6 | -2.5 | 0.78 |
| SAp-STE3 | 0.80 | 0.78 | -1.8 | -1.8 | 0.76 |
| SAp-STM1 | 1.29 | 1.26 | 2.2 | 1.9 | 0.68 |
| SAp-STM2 | 0.86 | 0.85 | -1.1 | -1.1 | 0.70 |
| SAp-STM3 | 1.04 | 0.95 | 0.4 | -0.3 | 0.66 |
| SAp-SEM1 | 1.17 | 1.17 | 1.4 | 1.2 | 0.68 |
| SAp-SEM2 | 0.79 | 0.76 | -1.9 | -2.0 | 0.80 |
| SAp-SEM3 | 0.78 | 0.76 | -1.9 | -2.1 | 0.79 |
| SAp-STEM1 | 0.71 | 0.79 | -2.6 | -1.8 | 0.77 |
| SAp-STEM2 | 0.75 | 0.73 | -2.2 | -2.3 | 0.77 |
| SAp-STEM3 | 0.98 | 1.00 | -0.1 | 0.0 | 0.69 |
| SAp-STEM4 | 0.61 | 0.65 | -3.7 | -3.1 | 0.76 |

## 3. Findings and Discussion

***STEM Application based on item logit value***

In this Wright map, the left side is item, while the right side is person map. This map is visualization of relationship of person and item in a single line (Abdullah, Noranee, & Khamis, 2017; W.J Boone et al., 2014). From the wright map, we can see that items are in the range of -1 to +1 logit value, while person spread widely from -3 to +6 logit value. Majority of person are in the range of -0.5 to +3 , indicating that they have higher average score compared to average of item logit value.

```
                          Item - MAP - Person
                             <rare>|<more>
 6                                 +  029SF  078BM
                                   |  017PM  130BF
 5                                 +
                                   |  027SF  033SM  093PM
 4                                 +  036SM  038CF  053SM  101CF
                                   |  024CM  120CM
                                   |  025CF  034SF
                                   |  135PM
 3                                 +  010PM  032PF  116CF
                                   |S 001CM  119PM
                                   |  002CM  097CF  102CF  112PF  136CF  137BM  139CF
                                   |  069BF  095CF  105CF  122SF
                                   |  009PF  048CM  085PF  086PF  115SF  117SM
 2                                 +  023SF  037CF  040BF  041SF  042BF  046SM
                                   |  012BM  081PM  094PF  100CF
                                   |  007BF  031PF  140PF
                                   |  028CM  087SF  104CF  121PM
                                   |M 006PF  011PF  016PM  035SF  059SF  060CF  075CF  077SF  088PF
                                      118PM  123SF  129SF  138CF
                                   |  056PF  127CM
 1                                T+  008BF  014CF  019CF  022CM  030CM  039PF  043CM  050SF  061BM
                                      070CF  072SF  073CM  103CM  106CF  132BM
        SAp-SEM2                   |  045SF  047CM  052SF  055SF  057BF  064BM  065BM  079BF  107CM
        SAp-SEM3  SAp-STEM1        |  080CF  083PF  126BF
        SAp-STE2  SAp-StEM2       S|  015PM  049SF  068BM  071SM  076SM  092PF  111CM  128CF
        SAp-STE3  SAp-STEM4 SAp-STM1 | 054PF 062CF  084PF  091PF  099CF
        SAp-SE2   SAp-STE1  SAp-STEM3 | 026SF 058PF 110SF  125CF
 0      SAp-SEM1  SAp-SM2         M+  013CF  018PF  044SF  063CF  066BM  074SF  089PF  098CF  124BF 131CM
        SAp-SE3  SAp-ST2 SAp-ST4 SAp-STM2 |S 003BF 051CF 096CF 113CF
        SAp-SM1                    |
        SAp-STM3                  S | 020CF  090PM  133CM
        SAp-SM3  SAp-ST1           |  005PM  114CM
        SAp-SE1                    |
-1      SAp-ST3                   T+
                                   |  021SF  108CM  134SF
                                   |T
-2                                 |  004BF
                                   |  082PF  109CM
-3                                 |  067BF
                             <frequ>|<less>
```

Figure 1. Wright Map

With regard to item logit, we also consider average and standard deviation of each sub domain in STEM application as detailed can be seen in Table 4. From the table, it is known that the hardest domain is SAp-SEM (M=0.533. SD=0.410), higher than SAp-STEM with 4 combination of specialization (M=0.270, SD=0.425). The results show how much Gen Z depends on technology, and the least domain is Science-Technology (M=-0.515, SD=-0.310), showing that pre-service science teachers can integrate technology better than integrating other specialization.

Table 4. Mean and Standard Deviation of STEM application

| Variable Name | Mean (M) | Standard Deviation (SD) |
|---|---|---|
| SAp-ST Application | -0.515 | -0.310 |
| SAp-SE Application | 0.364 | 0.538 |
| SAp-SM Application | -0.303 | 0.317 |
| SAp-STE Application | 0.286 | 0.155 |
| SAp-STM Application | -0.087 | 0.450 |
| SAp-SEM Application | 0.533 | 0.410 |
| SAp-STEM Application | 0.270 | 0.425 |

### ***STEM Application based on person logit value***

After analyzing this study using Winstep 3.73 Rasch model, we categorized STEM application into low (LVP ≤0.7), moderate (0.7<LVP<3.71) and high (LVP≥3.71). The categoration is based on person separation and the division by considering their range of logit value of person (LVP). Overall, most pre-service teachers (n=85 participants) are in moderate level of STEM application. In details, they are 44 participants (31.43%) in low category, 85 participants (60.71%) in moderate category and 11 participants (7.86%) in high category. The trend is closely similar when considering gender and specialization where most of participants are categorized in moderate level, the second one is low level, and the least one is high level of STEM application.

Table 5. STEM application

| | Low (LVP ≤0.7) | Moderate (0.7< LVP <3.71) | High (LVP≥3.71) |
|---|---|---|---|
| Overall | 44 | 85 | 11 |
| Gender | | | |
| Male | 13 | 29 | 6 |
| Female | 31 | 56 | 5 |
| Specialization | | | |
| Biology | 7 | 13 | 2 |
| Chemistry | 17 | 33 | 2 |
| Physics | 11 | 21 | 2 |
| Science | 9 | 18 | 5 |

### ***Based on Gender***

Gender is considerable factor in majority of social science research. In this study, comparison of STEM application based on gender is analyzed carefully using Rasch model with Winstep version 3.73. Each item comparison will consider Welch, Mantel Haenzel and Chi-Square, gender probability, which should be lower than 0.05 to be considered significantly different (Gocen & Sen, 2021; Rouquette, Hardouin, Vanhaesebrouck, Sébille, & Coste, 2019). Based on the analysis, it is found that gender do not influence STEM application in all items and all sub domain. It implies that Gen-Z can integrate Mathematics, Engineering and Technology in their online science teaching regardless their gender.

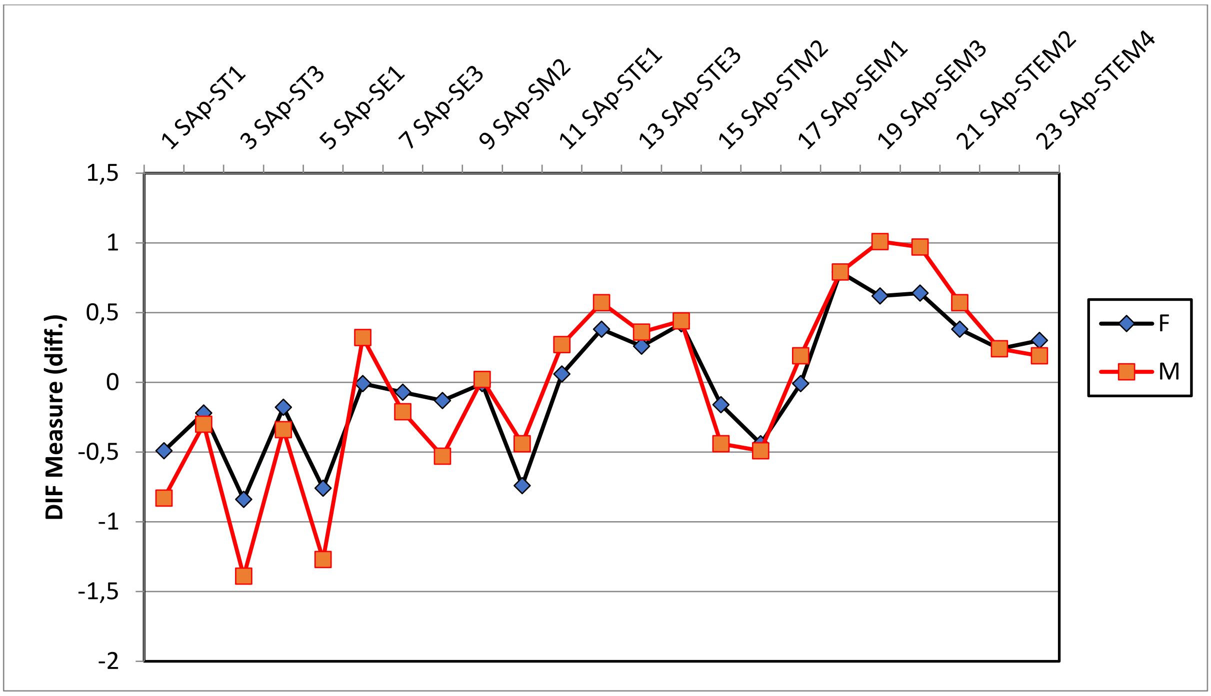


Figure 2. Comparison based on gender

***Based on Specialization***

Considering the result of careful analysis using Rasch model and taking into account the probability value of Welch, Mantel Haenzel and Chi-Square, in general, specialization does not influence Gen Z pre-service science teacher STEM application. However, some items showing any differences. The first item is SAp-ST4 “I use a ready-made technology tool (not made by myself) Application” between biology and science teachers.  Another item is “SAp-SE1 As a science teacher, in one term, I make learning media by myself” between Biology and science specialization as the previous item. The last item SAp-STM1” In my class, I usually use a technology tool to mathematically analyzing of data from observation (ex: use calculator, computer, mobile phone, ect)” which found any difference between Chemistry-Physics, Chemistry-Biology and Physics-Science.

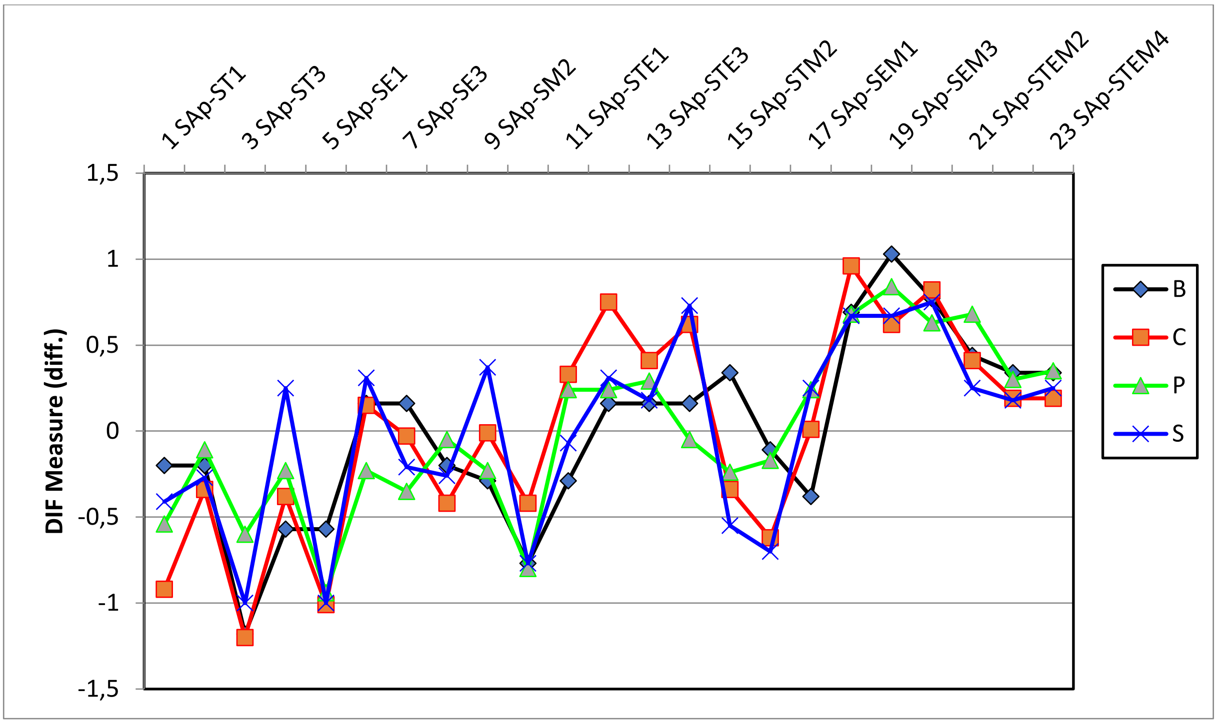


Figure 3. Comparison based on gender

## 4. Conclusion and Suggestion

### *Conclusion*

Consideing the research puroses and findings of the study, it is concluded that:

1) Most of GEN-Z pre-service science teachers applications are in moderate level
2) There is no significant influence of gender and specialization towards pre-service teachers STEM applications

### *Suggestion*

Findings of this studies can be the reference of indicating good future of STEM education if Gen Z pre-service teachers will be science teachers. In the future, it is vital to carry deeper analysis of predictors of their STEM applications, as well as the decent professional development to help them carry a better application of STEM education.

## References

Abdullah, N., Noranee, S., & Khamis, M. R. (2017). The use of rasch wright map in assessing conceptual understanding of electricity. *Social Sciences & Humanities*, *25*, 81–92.

Adams, D., Chuah, K. M., Sumintono, B., & Mohamed, A. (2021). Students ' readiness for e-learning during the COVID-19 pandemic in a South-East Asian university : a Rasch analysis pandemic. *Asian Education and Development Studies*. https://doi.org/10.1108/AEDS-05-2020-0100

Azman, M. N., Anom, M., Rashid, A., Rosnita, I., Faiq, M., & Aziz, A. (2021). Predicting Preferred Learning Styles on Teaching Approaches Among Gen Z Visual Learner.

*Turkish Journal of Computer and Mathematics Education*, *12*(9), 2969–2978.

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing Results From Empirical Research on Computer-Based Scaffolding in STEM Education: A Meta-Analysis. *Review of Educational Research*, *87*(2), 309–344. https://doi.org/10.3102/0034654316670999

Bogusevschi, D., Muntean, C. H., & Muntean, G. (2020). Teaching and Learning Physics using 3D Virtual Learning Environment : A Case Study of Combined Virtual Reality and Virtual Laboratory in Secondary. *Journal of Computers in Mathematics and Science Teaching*, *39*(1), 5–18.

Bond, T. ., & Fox, C. M. (2015). *Applying the Rasch model fundamental measurement in the human sciences* (3rd ed.). New York, NY, US: Routledge Taylor & Francis Group.

Boone, W.J, Staver, J. ., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Dordrecht: Springer Netherlands.

Boone, William J. (2016). Rasch analysis for instrument development: Why,when,and how? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

Chua, Y. P. (2013). *Mastering research statistics*. McGraw-Hill Education.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. *Educational Research* (Vol. 4). Cambridge: Pearson. https://doi.org/10.1017/CBO9781107415324.004

DeVellis, R. F. (2012). *Scale development: Theory and applications*. Newbury Park, CA: Sage Publications.

Elayyan, S. (2021). The future of education according to the fourth industrial revolution. *Journal of Educational Technology & Online Learning*, *4*(1), 23–30.

Fraenkel, J. R., & Wallen, N. E. (2009). *How to design and evaluate research in education. Boston: McGraw-Hill.* (7th ed.). New York, NY: McGraw-Hill.

Gocen, A., & Sen, S. (2021). A Validation of Servant Leadership Scale on Multinational Sample. *Psychological Reports*, *124*(2), 752–770. https://doi.org/10.1177/0033294120957246

Hidayat, R., Qudratuddarsi, H., Mazlan, N. H., & Zeki, M. Z. M. (2021). EVALUATION OF A TEST MEASURING MATHEMATICAL MODELLING COMPETENCY FOR INDONESIAN COLLEGE STUDENTS. *Journal of Nusantara Studies (JONUS)*, *6*(2), 133–155.

Iseppi, L., Rizzo, M., Gori, E., Nassivera, F., Bassi, I., & Scuderi, A. (2021). Rasch model for assessing propensity to entomophagy. *Sustainability*, *13*(8).

Iswadi, Syukri, M., Soewarno, Yulisman, H., & Nurina, C. I. E. (2020). A systematic literature review of science teachers' TPACK related to stem in developing a TPACK-stem scale. *Journal of Physics: Conference Series*, *1460*(1). https://doi.org/10.1088/1742-6596/1460/1/012105

Jin, Y., Rodriguez, C. A., Shah, L., & Rushton, G. T. (2020). Examining the Psychometric Properties of the Redox Concept Inventory: A Rasch Approach. *Journal of Chemical Education*, *97*(12), 4235–4244. https://doi.org/10.1021/acs.jchemed.0c00479

Khotimah, R. P., Adnan, M., Ahmad, C. N. C., & Murtiyasa, B. (2021). Science, Mathematics, Engineering, and Mathematics (STEM) Education in Indonesia: A Literature Review. *Journal of Physics: Conference Series*, *1776*(1). https://doi.org/10.1088/1742-6596/1776/1/012028

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement*, *3*(2), 103-122.

Manzoni, B., Caporarello, L., Cirulli, F., & Magni, F. (2021). The Preferred Learning Styles of Generation Z: Do They Differ from the Ones of Previous Generations?. *In Digital Transformation and Human Behavior (Pp. 55-67). Springer, Cham.*

Mustafa, N., Ismail, Z., Tasir, Z., & Mohamad Said, M. N. H. (2016). A meta-analysis on effective strategies for integrated STEM education. *Advanced Science Letters*, *22*(12), 4225–4288. https://doi.org/10.1166/asl.2016.8111

Poláková, P., & Klímová, B. (2019). Mobile technology and generation Z in the English language classroom – A preliminary study. *Education Sciences*, *9*(3), 1–11. https://doi.org/10.3390/educsci9030203

Qudratuddarsi, H., Sathasivam, R. V, & Hutkemri, H. (2019). Difficulties and correlation between phenomenon and reasoning tier of multiple-choice questions : A survey study. *Indonesian Research Journal in Education*, *3*(2), 249–264.

Romine, W. L., Schaffer, D. L., & Barrow, L. (2015). Development and application of a novel rasch-based methodology for evaluating multi-tiered assessment instruments: validation and utilization of an undergraduate diagnostic test of the water cycle. *International Journal of Science Education*, *37*(16), 2740–2768. https://doi.org/10.1080/09500693.2015.1105398

Rouquette, A., Hardouin, J. B., Vanhaesebrouck, A., Sébille, V., & Coste, J. (2019). Differential Item Functioning (DIF) in composite health measurement scale: Recommendations for characterizing DIF with meaningful consequences within the Rasch model framework. *PLoS ONE*, *14*(4), 1–16. https://doi.org/10.1371/journal.pone.0215073

Sadhu, S., & Laksono, E. W. (2018). Development and validation of an integrated assessment for measuring critical thinking and chemical literacy in chemical equilibrium. *International Journal of Instruction*, *11*(3), 557–572. https://doi.org/10.12973/iji.2018.11338a

Shultz, K. S., Whitney, D. J., & Zickar, M. J. (2014). *Measurement theory in action: Case studies and exercises* (2nd ed.). Third Avenue, New York: Routledge/Taylor & Francis Group.

Sukendro, S., Habibi, A., Khaeruddin, K., Indrayana, B., Syahruddin, S., Makadada, F. A., & Hakim, H. (2020). Using an extended Technology Acceptance Model to understand students' use of e-learning during Covid-19: Indonesian sport science education context. *Heliyon*, *6*(11), e05410. https://doi.org/10.1016/j.heliyon.2020.e05410

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan*. Cimahi: Trim Komunikata.

Suryadi, B., Hayat, B., Dwirifqi, M., & Putra, K. (2021). The Indonesian version of the Life Orientation Test-Revised ( LOT-R ): Psychometric properties based on the Rasch model The Indonesian version of the Life Orientation Test-Revised ( LOT-R ): Psychometric properties based on the Rasch model. *Cogent Psychology*, *8*(1). https://doi.org/10.1080/23311908.2020.1869375

Szymkowiak, A., Melović, B., Dabić, M., Jeganathan, K., & Kundi, G. S. (2021). Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people. *Technology in Society*, *65*(March). https://doi.org/10.1016/j.techsoc.2021.101565

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology :

What is it and why use it ? When should it be applied , and what should one look for in a Rasch paper ? *Arthritis & Rheumatism*, *57*(8), 1358–1362. https://doi.org/10.1002/art.23108

Wind, S. A., & Gale, J. D. (2015). Diagnostic opportunities using rasch measurement in the context of a misconceptions-based physical science assessment. *Science Education*, *99*(4), 721–741. https://doi.org/10.1002/sce.21172

Yahya, M. S., & Hashim, H. (2021). Interdisciplinary Learning and Multiple Learning Approaches in Enhancing the Learning of ESL among STEM Learners. *Creative Education*, *12*(05), 1057–1065. https://doi.org/10.4236/ce.2021.125078

Yildirim, B. (2016). An analyses and meta-synthesis of research on STEM education. *Journal of Education and Practice*, *7*(34), 23–33.

Zahidi, A. M., Ing, O. S., Yusof, R., Kanapathy, S. A., Ismail, M. J., Huay, & You, W. (2021). Effect of science camp for enhancing STEM skills of gifted young scientists. *Journal for the Education of Gifted Young Scientists*, *9*(1), 15–26. https://doi.org/10.17478/jegys.821700

Zuidgeest, M., Hendriks, M., Koopman, L., Spreeuwenberg, P., & Rademakers, J. (2011). Comparison of a postal survey and mixed-mode survey using a questionnaire on patients' experiences with breast care. *Journal of Medical Internet Research*, *13*(3).