

Psychometric Assessment of the Short Grit Scale (Grit-S) using the Rasch Model in a Malaysian Polytechnic Students' Sample

Mohd Effendi a.k.a. Ewan Mohd Matore^{1*}, Norazlinda Mohamad²,
Ahmad Zamri Khairani³, Tanesh N. Segar¹

**Corresponding Author*

^{1*} Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia

² Europasia Engineering Services Sdn. Bhd, Petaling Jaya, Malaysia

³ School of Educational Studies, Universiti Sains Malaysia, Penang, Malaysia

** Corresponding Author*

effendi@ukm.edu.my, linda@europasia.com.my, ahmadzamri@usm.my, p121083@siswa.ukm.edu.my
Tel: +60192896080

Abstract

The study investigated the psychometric assessment of grit questions in the setting of a technical student using the Rasch model. Random sampling techniques were chosen among 468 polytechnic students. The results showed eight Short Grit Scale (Grit-S) items met the requirements of the Rasch model. These items can measure grit for technical students' context, but there is ample room for further progress in determining the quality of grit items' development in different settings. In a nutshell, Grit-S needs to generate more items to cater to people with higher abilities so that they can be tested with more complex items.

Keywords: grit; psychometric assessment; technical students; Rasch model

eISSN: 2398-4287 © 2023. The Authors. Published for AMER and cE-Bs by e-International Publishing House, Ltd., UK. This is an open-access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). Peer-review under the responsibility of AMER (Association of Malaysian Environment-Behaviour Researchers), and cE-Bs (Centre for Environment-Behaviour Studies), College of Built Environment, Universiti Teknologi MARA, Malaysia
DOI: <https://doi.org/10.21834/e-bpj.v8iS115.5088>

1.0 Introduction

In the new global economy, grit has emerged as a significant issue for human adaptability. Duckworth (2018) defines grit as "long-term dedication and tenacity towards meaningful aims." When faced with challenges at work, technical professionals require grit, and having things capable of determining grit concerns is vital. As defined by Stoffel and Cain (2018), grit is the ability to keep going in the face of hardship. Grit is required for technical employees to tackle life's difficulties and to teach us about passion. Polytechnic students have the fortitude needed to overcome hurdles in the workplace. Little is known about their grit, and it is questionable whether grit is so closely tied to polytechnic students. As a result, more precise grit measuring items are required for technical students to provide recommendations on how to be better workers. The Short Grit Scale or Grit-S is a helpful scale for determining grit (A. L. Duckworth & Quinn, 2009). Previous international research from Malaysia (González-Bernal et al., 2022; Postigo et al., 2023) has only dealt with the psychometrics of Grit-S using classical test theory compared to the modern approach for Rasch.

Aarepattamanni and Khine (2018) used Rasch analysis among 777 adolescents hailing from Arab states of the Persian Gulf. Unfortunately, they did not take technical student respondents as a sample. The study by Segar and Mohd Matore (2023) purposefully increased the ambiguity of grit level measurement in Mathematics exclusively in the setting of Tamil National Type Primary School (SJKT) students. This study aims to use Rasch analysis to assess the psychometric features of grit items in the context of technical students. They possessed a critical source of trained and technically capable individuals to help the economy. Moreover, they are meeting the demands of the workforce, addressing skill gaps, and encouraging innovation and entrepreneurship. This empirical evidence will enable us to move forward to improve the lacking issues of measurement aspect of grit among Malaysian technical students.

2.0 Literature Review

2.1 Definition of grit

Grit is a characteristic of personality that describes a person's dedication and enthusiasm for long-term objectives and the ability to maintain effort and interest for lengthy periods in the face of difficulties, challenges, and failures. Students should be provided with long-term opportunities to accomplish work since grit can be built and taught (Tiandem-Adamou & Hargis, 2022). Cultivate courage in pupils so they can endure, even in the face of hardship, until the goal is reached (Yazon et al., 2021). Several past grit studies have been conducted concerning assessing psychometric characteristics (Ezaty et al., 2019; González-Bernal et al., 2022; Postigo et al., 2023). However, their research focuses more on psychometric assessment using classical test theory than modern theories such as the Rasch model. Previous literature established the idea of conducting a psychometric assessment of grit items using the Rasch model and is appropriate in the context of this research framework because this measurement model has the potential to overcome limitations in classical test theory.

2.2 Short Grit Scale (Grit-S)

In 2007, Angela Duckworth and her colleagues created the Short Grit Scale (Grit-S) as an improved version of the previous Grit Measure (Duckworth & Quinn, 2009). The Grit-S consists of eight items scored on a Likert scale of one to five, with answers varying from "not at all like me" to "very much like me." The items measure two aspects of grit: persistence of effort and consistency of interest. Table 1 displays eight Grit-S items:

Table 1. Short Grit Scale (Grit-S)

No of items	Items of Grit-S
1	I have overcome setbacks to conquer a significant challenge.
2	New ideas and projects sometimes distract me from previous ones.
3	I have been obsessed with a specific idea or project for a short time but later lost interest.
4	My interests change from year to year.
5	Setbacks do not discourage me.
6	I have been working on the same thing for years.
7	I am a hard worker.
8	I often set a goal but later choose to pursue a different one.

(Source:) Duckworth & Quinn (2009)

3.0 Methodology

The survey research design was adopted, and the research approach was purely quantitative. The survey was chosen because this design is suitable for research involving many respondents. This study comprised 468 polytechnic students across five geographical regions (West, North, East, South, and Borneo) using a clustered multistage stratified proportionate sample with 10% for each stratum, as suggested by Gay and Mills (2018). The classifications are program type, academic year, and gender. The selection of this stratum aims to ensure that the respondents can be reached well across various layers of sample characteristics. This study has limitations in that it does not involve students undergoing industrial training and the final semester of studies to avoid disrupting students. The data collection procedure involves requesting permission from the Ministry of Higher Education (MOHE) and polytechnics. It was done through student affairs officers appointed at the institution to avoid researcher bias.

The eight Grit-S items were used to generate the items (Duckworth & Quinn, 2009). Consistency of Interest and Perseverance of Effort are first-order latent variables influencing grit, a second-order latent factor. Items 2, 4, 7, and 8 are scored with the following number of points on a scale based on five Likert points: (5 = Very much like me, 4 = Mostly like me, 3 = Somewhat like me, 2 = Not much like me, and 1 = Not at all). Assign the following points to questions 1, 3, 5, and 6, and vice versa. The entire score will be added jointly and multiplied by 8. This multiplication is sufficient to get the mean score stated in the original instrument. The maximum score on this scale is 5 (extremely gritty), and the lowest is 1 (not at all gritty). The item's psychometrics will be assessed using the Rasch model, such as analysis of item fit, unidimensionality, local independence, gender differential item functioning (GDIF), Wright map, reliability, and separation index.

4.0 Findings and Discussions

4.1 Item fit

The Mean Square Fit Statistic (MNSQ) with the Z standard (Zstd) will determine the Rasch model's first assumption, item fit. The degree to which individual test items or questions comply with the model's expectations is item fit. Item fit, in particular, relates to how well an item fits inside the model's overarching hierarchy of item difficulty or severity. Table 2 and Figure 1 demonstrate that the results revealed eight items in the 0.6-1.4 logit range with MNSQ and 2.0 with Zstd, as Bond et al. (2021) recommended. The MNSQ in Table 2 shows the range for infit was 0.7 to 1.19 logit, and the MNSQ range for the outfit was 0.69 to 1.23 logit. Figure 1 demonstrates that G6 and G8 have infit Zstd values greater than 2.0 and should be reviewed. Bond et al. (2021) proposed that the polarity of the item be stated as positive and greater than 0.4. The polarity item ranges from 0.52 to 0.68 logit, as shown in Table 2.

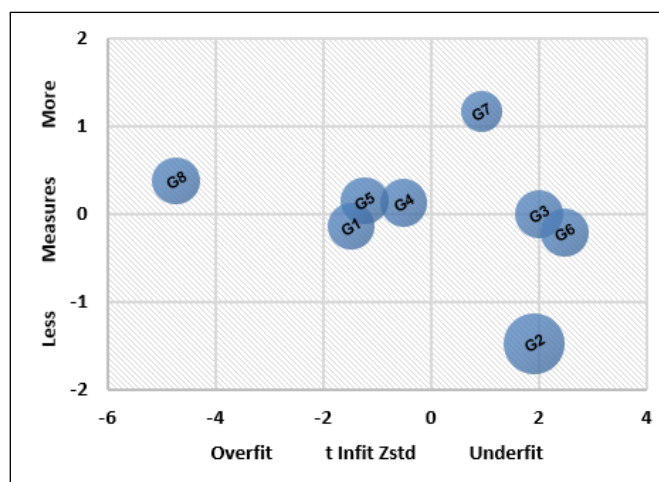


Fig. 1: Infit Bubble chart

No item	Total score	Login	S. E	MNSQ		PTMEA	
				Infit	Outfit	Corr.	Exp.
3	1805	0.00	0.07	1.15	1.23	0.60	0.60
6	1845	-0.21	0.07	1.19	1.18	0.58	0.59
2	2047	-1.47	0.09	1.14	1.10	0.52	0.52
7	1541	1.16	0.06	1.06	1.09	0.68	0.66
4	1778	0.13	0.07	0.96	0.97	0.62	0.61
1	1831	-0.14	0.07	0.90	0.94	0.52	0.59
5	1774	0.15	0.07	0.91	0.89	0.66	0.61
8	1727	0.38	0.07	0.70	0.69	0.63	0.62

4.2 Unidimensionality

Based on the unidimensionality, things will be measuring only one construct. Dimensionality is defined as the ability to determine an instrument in only one direction. Table 3 demonstrates that 44.1% of the natural variance explained by measurements fulfills the criterion by at least 40% (Bond et al., 2021). In the first comparison, the unexplained variation was 11.6%. (below the criterion of 15%) (Mofreh, 2019), while the Eigenvalue was 1.7 (below the requirement of 2 as the accepted value) (Bond et al., 2021).

		Empirical		Modeled
Total raw variance in observations		14.3	100.0%	100.0%
Raw variance explained by measures		6.3	44.1%	43.9%
Raw variance explained by persons		3.2	22.3%	22.2%
Raw variance explained by items		3.1	21.8%	21.7%
Raw unexplained variance (total)		8.0	55.9%	100.0%
Unexplained variance in first contrast		1.7	11.6%	20.8%

4.3 Local independence

The third premise of the Rasch model is local independence. There should be no association between residual items in Table 4 for possible pairing items. Table 4 shows that the greatest combination is ($r = -0.32$), and the value is less than 0.5 and close to Abdellatif (2023) earlier study. The range of correlation of the local independence is -0.18 to -0.32.

Correlation	Item number	Item number
-0.32	G2	G7
-0.27	G3	G7
-0.26	G4	G8
-0.25	G3	G6
-0.24	G2	G8
-0.22	G4	G7
-0.20	G6	G7
-0.20	G6	G8
-0.19	G5	G7
-0.18	G3	G5

4.4 Gender differential item functioning (GDIF)

Gender Differential Item Functioning, or GDIF, was a method of statistics used in psychometrics to determine if test items or questions act differently when administered to males and females. The GDIF examines the accomplishments of different categories of participants on specific questions while adjusting for their general ability or characteristic level.

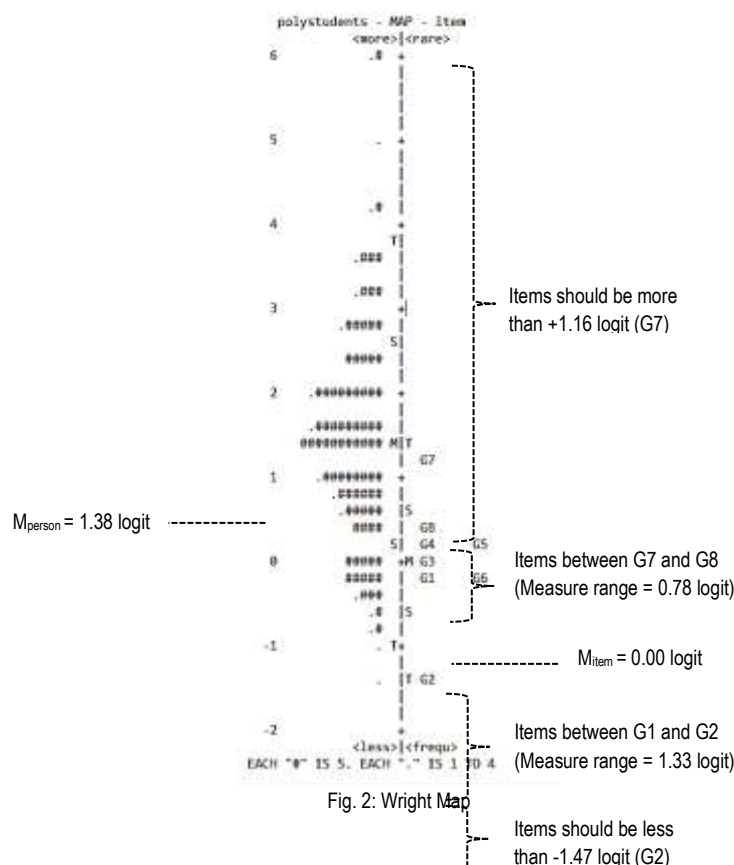
Table 5 reveals that, except for G5, all eight factors match the GDIF criteria. G5 should be 0.5 on the DIF contrast, 2.0 on the t value, and p larger than 0.05. (Sovey et al., 2022). DIF Contrast ranges from -0.34 to 0.32, while t ranges from 1.86 to -2.38. G5 had a T value of -2.38, which was outside range. Both tests, Welch's t-test, and Mantel-Haenszel, will be used to satisfy the t-test criteria. A Welch's t-test was also utilized to figure out the difference, in addition to a Mantel and Hanzel analysis to contrast observed with anticipated DIF values. Item G5 did not meet the $p > 0.05$ requirement and recorded the probability as (Welch = 0.0180; Mantel-Haenszel = 0.0113). This signifies that G5 passed the DIF contrast but failed the t value, $p > 0.05$. G5 (Setbacks do not discourage me) is likewise biased against male pupils and should be improved in item phrase construction.

Table 5. Gender differential item functioning (GDIF)

No item	Total score	Login	DIF Measure (Male)	DIF Measure (Female)	S. E	DIF Contrast (DIF Size)	t	Probability (Welch)	Probability (Mantel-Haenszel)
G1	1831	-0.14	-0.24	-0.06	0.07	-0.18	-1.21	0.2275	0.1985
G2	2047	-1.47	-1.30	-1.62	0.09	0.32	1.86	0.0635	0.0602
G3	1805	0.00	0.10	-0.09	0.07	0.19	1.30	0.1931	0.3653
G4	1778	0.13	0.13	0.13	0.07	0.00	0.00	1.0000	0.9749
G5	1774	0.15	-0.04	0.30	0.07	-0.34	-2.38	0.0180	0.0113
G6	1845	-0.21	-0.14	-0.27	0.07	0.13	0.89	0.3755	0.4198
G7	1541	1.16	1.10	1.21	0.06	-0.11	-0.84	0.3992	0.3989
G8	1727	0.38	0.43	0.33	0.07	0.09	0.69	0.4902	0.2090

Research suggests that males and females may respond differently to setbacks, but these differences are not due to biological or innate factors. Instead, they are likely the result of societal and cultural expectations and stereotypes. Encouragement should not be gender-specific and should be provided equally to both males and females. Everyone, regardless of gender, benefits from encouragement and support. Unfortunately, cultural stereotypes and biases often lead to different expectations and treatment of males and females. Encouragement can take many forms, such as positive feedback, constructive criticism, mentorship, and emotional support. By providing equal encouragement to all individuals, regardless of gender, we can help to create a more supportive and inclusive society.

4.5 Wright Map



G7 expressing "I am a hard worker" (measure = +1.16 logit) is probably the most challenging item for students to endorse, as seen in Figure 2, while G2 (measure = -1.47 logit) is the easiest. The logit range spread that satisfied the item difficulty was 2.63. It implies that more items should be developed to test the respondents with higher-ability respondents. The logit of the objects should be more than +1.16 logit. Other spaces that new grit items can fill are with difficulty more than G7 (Measure range = more than 1.16 logits), between G7 and G8 (Measure range = 0.78 logits), between G1 and G2 (Measure range = 1.33 logits) and below than G2 (Measure range = less than -1.47 logits),

4.6 Reliability and separation index

Tables 6 and 7 present a summary of statistics regarding the item with individual. The item's reliability was 0.99, but the person's reliability was 0.72. Item reliability refers to the consistency and stability of individual test items or questions over time or across different administrations. A highly reliable item gets an identical response from test-takers with the same grasp of the trait or ability evaluated. Person reliability, on the other hand, refers to the consistency and stability of individual test-takers' scores over time or across different administrations. A highly reliable test consistently measures the same trait or ability across the different individual's traits and situations.

The person separation value is 1.60, and the item separation value is 9.14. All values are suitable and sufficient. Even if the person separation index is less than 2, the separation index can be accepted if more than 1.5 (Post et al., 2022). A higher separation index for items indicates that the test or scale is more precise and can better differentiate between individuals with different levels of trait or ability. It means there may be some overlap in scores between individuals with different levels of the trait or ability, making it more difficult to differentiate between them accurately.

Table 6. Summary statistic for the person (N = 459; extreme score = 9)

	Raw Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	30.5	8.0	1.38	0.56	1.01	-0.10	1.01	-0.1
Standard Deviation	4.0	0.0	1.20	0.09	0.78	1.30	0.78	1.3
Max	39.0	8.0	5.10	1.08	5.88	5.40	5.73	5.3
Min	18.0	8.0	-1.41	0.42	0.03	-4.00	0.04	-3.9
Real RMSE	0.64	True SD	1.02	Separation	1.60	Person Reliability		0.72
Model RMSE	0.56	True SD	1.06	Separation	1.89	Person Reliability		0.78

Person Raw Score-To-Measure Correlation = .97

Cronbach Alpha (KR-20) Person Raw Score Reliability = .77

Table 7. Summary statistic for the item (8 items)

	Raw Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	1793.5	468.0	0.00	0.07	1.00	-0.1	1.01	0.0
Standard Deviation	131.0	0.0	0.69	0.01	0.15	2.3	0.16	2.5
Max	2047.0	468.0	1.16	0.09	1.19	2.5	1.23	3.1
Min	1541.0	468.0	-1.47	0.06	0.70	-4.7	0.69	-5.0
Real RMSE	0.7	True SD	0.68	Separation	9.14	Item Reliability		0.99
Model RMSE	0.7	True SD	0.68	Separation	9.47	Item Reliability		0.99

UMean = 0.000 UScale = 1.000

Item Raw Score-To-Measure Correlation = -0.99

Cronbach Alpha was calculated as 0.77 in Table 6 in order to test internal consistency. It evaluates how well a group of test items or questions measures a particular underlying construct or characteristic. Cronbach's alpha at 0.7 or more is considered satisfactory (Hair et al., 2020). If Cronbach's alpha is elevated, it suggests that the test or scale items consistently and reliably measure the identical underlying construct or feature.

As mentioned in the literature review, the findings prove that the eight Short Grit Scale (Grit-S) items meet the psychometric assessment among technical students using the Rasch model. The items were found to have good validity and reliability considering the items and individuals. It is interesting to note that most items met the assumptions of modern test theory in this study using the Rasch model and also meeting classical test theory as in previous studies (Ezaty et al., 2019; González-Bernal et al., 2022; Postigo et al., 2023). Surprisingly, few differences were found in the item assessment for both measurement theories, even though the nature was different. This study produced results that verify the findings of a great deal from the previous work in this field. This result may be explained by the fact that the item's quality is still consistent even though this research was conducted on Malaysian local polytechnic students.

5.0 Conclusion

In summary, these results show that items fulfilled the Rasch model assumptions. Except for G5 (Setbacks do not discourage me), the results show that most of the eight Short Grit Scale (Grit-S) items meet the Rasch model analysis. However, this assessment of analysis has several limitations. Firstly, the psychometric assessment only looks at items and not people. The assessor must accommodate the modification of the person's mismatch to increase the value by screening the person. Indirectly, the item psychometric assessment for Grit-S using the Rasch model will provide compelling empirical evidence to improve validity and reliability. Secondly, the research context only focuses on polytechnic students. A much more systematic approach would be better if respondents could cater to a broader context

of technical students such as Pusat Latihan Teknologi Tinggi (ADTEC), Institut Teknikal Jepun Malaysia (JMTI), Institut Latihan Perindustrian (ILP) or Institut Kemahiran MARA (IKM). The present study provides additional evidence contribution concerning the measurement theory applicability and enhances our understanding of the body of knowledge on grit. Further research might explore the potential of the grit items with new or different constructs in our local context. It is recommended that Malaysia needs its grit instrument to measure our local students accurately for identification purposes. Students with weak grit can be appropriately guided via intervention.

Acknowledgments

The Faculty of Education at UKM approved the project under the GG-2022-020 (From Research Fund of FPEND) and GP-2021-K021854 (Publication Incentive Grant) grants. Special thanks also to University Research Group (KPU), Educational Evaluation UKM.

Paper Contribution to Related Field of Study

This research extends our knowledge to the aspect of performance measurement. Assessment of grit measurement with high validity and reliability using the Rasch model contributed to providing empirical evidence from technical students' context with the applications of modern measurement theory such as the Rasch model. The Rasch model is essential because it provides a more accurate estimate of an individual's ability than traditional methods (classical test theory), which only provide a raw score based on the number of items answered correctly.

References

- Aarepattamannil, S., & Khine, M. S. (2018). Evaluating the Psychometric Properties of the Original Grit Scale Using Rasch Analysis in an Arab Adolescent Sample. *Journal of Psychoeducational Assessment*, 36(8), 856–862. <https://doi.org/10.1177/0734282917719976>
- Abdellatif, H. (2023). Test results with and without blueprinting: Psychometric analysis using the Rasch model. *Educacion Medica*, 24(3), 100802. <https://doi.org/10.1016/j.edumed.2023.100802>
- Bond, T. G., Yan, Z., & Heene, M. (2021). Applying the Rasch Model: Fundamental Measurement in the Human Sciences. In *Applying the Rasch Model* (4th ed.). Routledge. <https://doi.org/10.4324/9781410614575>
- Duckworth, A. (2018). *GRIT: The power of passion and perseverance*. Simon and Schuster.
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91(2), 166–174. <https://doi.org/10.1080/00223890802634290>
- Ezaty, H., Hasan, A., Hanum, J., & Khaiyom, A. (2019). A Study Protocol on the Adaptation and Validation of the Malay Version of the Short Grit Scale (Grit-S) on Malaysian University Students. *Malaysian Journal of Psychiatry*, 28(2), 1–9.
- Gay, L. R., & Mills, G. E. (2018). *Educational research: competencies for analysis and applications* (12th ed.). Merrill Prentice Hall.
- González-Bernal, J., Gonzalez-Bernal, S., Salavera, C., Fernández-Ortega, C., Trigueros Ramos, R., Aguilar-Parra, J. M., & González-Santos, M. J. (2022). Adaptation and Testing of the Factorial Structure of the Physical Education Grit Scale for Use in Secondary Education in Spain. *International Journal of Environmental Research and Public Health*, 19(16). <https://doi.org/10.3390/ijerph191610008>
- Hair, J. F., Celsi, M. W., & Harrison, D. E. (2020). *Essentials of marketing research* (5th ed.). McGraw-Hill Education.
- Post, M. W. M., Fellinghuier, C. S., Charlifue, S., New, P. W., Forchheimer, M. B., & Tate, D. G. (2022). Rasch Analysis of the International Quality of Life Basic Data Set Version 2.0. *Archives of Physical Medicine and Rehabilitation*, 103(11), 2120–2130. <https://doi.org/10.1016/j.apmr.2022.02.018>
- Postigo, A., Barria, J., Cuesta, M., & García-Cueto, E. (2023). Psychometric Properties of the Chilean Version of the Oviedo Grit Scale Collabra : Psychology. *Collabra: Psychology*, 9(1), 1–14. <https://doi.org/https://doi.org/10.1525/collabra.57516>
- Segar, T. N., & Mohd Matore, M. E. E. (2023). Fantastic GRIT : Celebrating the determination of Tamil National Type Primary School (SJKT) Students in Mathematics Learning Fantastic GRIT : Celebrating the determination of Tamil National Type Primary School (SJKT) Students in Mathematics Learning. *International Journal of Academic Research in Business and Social Sciences*, 13(3), 1609–1621. <https://doi.org/10.6007/IJARBS/v13-i3/16311>
- Sovey, S., Osman, K., & Mohd Matore, M. E. E. (2022). Gender differential item functioning analysis in measuring computational thinking disposition among secondary school students. *Frontiers in Psychiatry*, 13(November), pp. 1–14. <https://doi.org/10.3389/fpsy.2022.1022304>
- Stoffel, J. M., & Cain, J. (2018). Review of grit and resilience literature within health professions education. *American Journal of Pharmaceutical Education*, 82(2), 124–134. <https://doi.org/10.5688/ajpe6150>
- Tiandem-Adamou, Y. M., & Hargis, J. (2022). Grit and Chinese Students' Foreign Language Enjoyment in University. *International Journal on New Trends in Education and Their Implications (IJONTE)*, 13(2), 76–89. <https://doi.org/10.21608/psjh.2022.250026>
- Yazon, A. D., Ang-Manaig, K., & Adrian, T. W. C. (2021). A Correlational Study on Mindset, Grit, and Adversity Quotient of Pre-Service Teachers: Evidence in the Philippines and Hong Kong. *International Journal of Management, Entrepreneurship, Social Science and Humanities*, 4(2), 174–181. <https://doi.org/10.31098/ijmesh.v4i2.784>