Studies in Philosophy of Science and Education
Homepage: https://scie-journal.com/index.php/SiPoSE
Email: sipose@scie-journal.com

p-ISSN: 2722-3523; e-ISSN: 2721-9267
SiPoSE, Vol. 5, No. 1, March 2024
Page: 41-47
©2024 SiPoSE

Item Test Analysis of Entrepreneurship Readiness Instrument for Vocational School in East Java Indonesia

***S R Hakim[1,2], Ekohariadi[1], Suparji[1], Harmanto[1], E Wiwin[3]**
[1]Postgraduate School, Universitas Negeri Surabaya, Indonesia
[2]Universitas Negeri Yogyakarta, Indonesia
[3]Cabdin Pendidikan Wilayah Sidoarjo, Indonesia

Article Info

Article history:

Received August 21, 2023
Revised March 20, 2024
Accepted March 29, 2024
Available Online March 31, 2024

Keywords:

Educational levels;
Entrepreneurship readiness;
Indonesia labor;
Open unemployment rates;
Vocational High School;

ABSTRACT

Vocational High Schools (SMK), recognized as a solution to address unemployment in Indonesia, simultaneously face the challenge of having the highest open unemployment rate (TPT) compared to other educational levels, standing at 10.38%. Enhancing the quality of questions assessing entrepreneurial readiness is a pivotal approach in tackling this issue. Various factors contribute to students' preparedness for entrepreneurship. This study aims to identify these components by surveying 176 respondents. Employing the Entrepreneurial Readiness Test and analyzing it through Classical Test Theory (CTT) and the Rasch Analysis Model, the research revealed that entrepreneurial mental attitude, risk-taking ability, creativity, motivation, and entrepreneurial skills significantly influence students' readiness for entrepreneurship. The analysis yielded IMS (Infit mean-square) scores ranging from 0.88 to 1.17 and OMS (Outfit mean-square) values between 0.86 to 1.35, indicating a fit for all 50 items in the entrepreneurial readiness test.

https://doi.org/10.46627/sipose

INTRODUCTION

Improving the quality of education remains a persistent concern for the Indonesian government, being a key focus in national educational development initiatives. Enhancing education is vital for bolstering the country's human resources. The notable challenges include the high unemployment rate among graduates of vocational schools, inconsistent policies, insufficient teacher competencies, and inadequate support systems, as highlighted by critics (Suharno, 2018). Industries employing these graduates demand academic competencies and practical skills (Russell, 2018). Attaining high-quality human resources necessitates high-quality education. However, Indonesia faces numerous hurdles in this endeavor, encompassing the quality of educators and educational staff, student capabilities, infrastructure, and learning processes, including evaluation and assessment methods.

In Indonesia, there persists a need for significant enhancements in both the quantity and caliber of vocational education. The implementation of technology and vocational education encounters various challenges. Suranto (2006) highlights the issue of low community involvement in financing education, particularly in engineering, vocational education, and occupational training, despite a decrease in interest in these fields among college applicants. Moreover, governmental management of vocational schools and post-graduation support remains a pressing concern. Challenges encompass the low quality and relevance of graduates and the scarcity of qualified teachers for vocational institutions. Despite a relatively high employment rate (85%), many industries voice dissatisfaction with the preparedness of vocational graduates for the workforce (Suharno, 2018). Interviews with six company managers

https://doi.org/10.46627/sipose.v5i1.330

revealed that vocational graduates often lack essential skills, falling below the expected proficiency level (Russell, 2018). Consequently, while these graduates are sought after by employers, their perceived inadequacy remains a persistent issue.

According to the UNDP's 2018 report, Indonesia's standing has declined. Indonesia's Human Development Index (HDI) stands at 0.694, ranking 116th globally, trailing behind Malaysia (57), Singapore (9), and Thailand (83). Data indicates a diminishing competitiveness of Indonesian workers compared to their counterparts in ASEAN countries. This decline correlates with prevailing labor conditions. Presently, Indonesia grapples with persistently high levels of unemployment. As per the Central Bureau of Statistics (refer to figure 1), the national open unemployment rate stood at 10.32% of the workforce in February 2022. The elevated unemployment rate may stem from either the inadequate competencies of Indonesian human resources or a lack of employment opportunities to accommodate the influx of graduates from educational institutions.

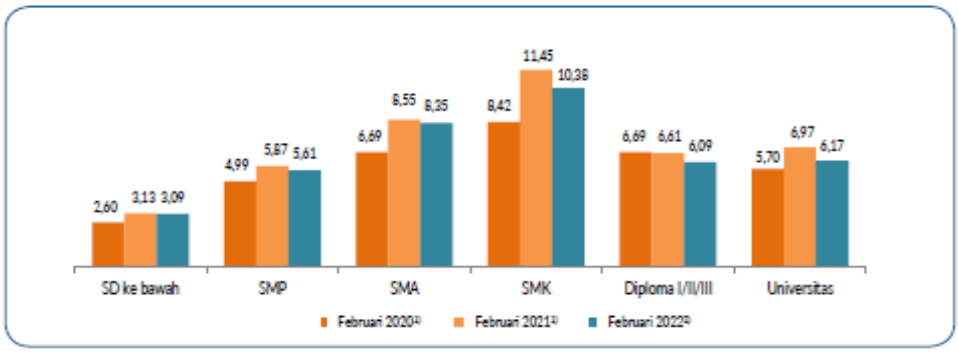


Figure 1. Indonesia labor condition

Based on this data, numerous factors contribute to the extent of open unemployment rates. One such factor is the substandard quality of questions used to gauge entrepreneurial readiness in learning assessment procedures. Assessment activities hold considerable importance in determining the success of educational endeavors (Mansyur, 2015). Measurement, a pivotal component, precedes all evaluation actions and calculates students' proficiency in a given subject (Mardapi, 2012). Hence, the issue of measurement quality significantly impacts the evaluation process.

Typically, test instruments serve the purpose of evaluating student achievement. A test is a tool to capture a snapshot of an individual's behavior (Allen, 1979). Similarly, Mardapi (2012) echoes this sentiment, defining a test as a means of measurement to gather information about an object's characteristics. When constructing tests, attention is given to the characteristics of the items comprising them. These characteristics ensure a balanced mix of item difficulty levels—ranging from easy to challenging—about the subject matter being assessed. Additionally, an essential aspect involves the differential power of items in discerning between high-ability and low-ability student groups, thereby facilitating a more accurate depiction of student learning outcomes. Assessing item characteristics can be accomplished through two approaches: Classical Test Theory (CTT) and Item Response Theory (IRT).

RESEARCH METHOD

The author designed a test to investigate the utilization of the Entrepreneurial Readiness Test in courses focusing on Creative Products and Entrepreneurship. This test comprises 50 multiple-choice items to assess entrepreneurial mental attitude, risk-taking ability, creativity, motivation, and entrepreneurial skills. Student responses to these 50 test items underwent analysis through computer programs, specifically the Test Analyze Program (Brooks, 2003) and ConQest (Wu et al., 1998). The Test Analyze Program estimates the item difficulty parameter (p) utilizing classical test theory, while ConQest is used to compute the item difficulty parameter (β) employing item response theory.

https://doi.org/10.46627/sipose.v5i1.330

A measure of entrepreneurial readiness underwent analysis, with the test results examined using the item response theory within the Rasch model, employing STATA 16 software. Fit analysis was conducted to assess the unidimensionality of the entrepreneurial readiness test items, which is a fundamental assumption in item response theory. Unidimensionality implies that the test items measure a single ability (Hulin et al., 1983). One method of evaluating unidimensionality is examining the fit of the data to the Rasch model (Smith, 2001). Mean-square infit (IMS) and mean-square outfit (OMS) statistics were utilized to determine whether the Rasch model could predict each respondent's response accurately (Smith, 2001). IMS and OMS statistics indicate the degree of agreement between observed data and model-predicted values. The selection of test items in this study was guided by the Rasch model, with items chosen based on IMS and OMS values. Tables were constructed to interpret the significance of IMS and OMS values (Linacre, 2002b). Table 1 illustrates the interpretations of IMS and OMS values. The test items exhibited IMS and OMS values ranging from 0.5 to 1.5.

Table 1. Range of IMS and OMS values

Mean-square Value	Implication for Measurements
>2.0	Damaging the measurement system
1.5 - 2.0	Has no meaning for measurement
0.5 - 1.5	Useful for measurement
<0.5	Not useful for measurement but not harmful

RESULTS AND DISCUSSION

Results

The test comprises 50 things of choice reactions utilizing 176 professional school understudies with computer ability in East Java, Indonesia. Responses to the taking after 50 test focuses were analyzed using the Victory computer program. Table 2 shows that the assessment comes from the chosen reaction sort parameters. There are 50-grain trouble parameters (β).

Table 2. Test item difficulty parameters

No.	Par β	Difficulty	No.	Par β	Difficulty
1	-0.418	0.5568	26	-0.567	0.5909
2	-0.149	0.4943	27	-1.043	0.6932
3	-0.932	0.6705	28	1.003	0.2500
4	-0.027	0.4659	29	1.035	0.2443
5	-1.578	0.7898	30	1.315	0.1989
6	-1.158	0.7159	31	-2.075	0.8580
7	-0.319	0.5341	32	1.314	0.1989
8	-0.987	0.6818	33	-0.773	0.6364
9	0.298	0.3920	34	0.297	0.3920
10	0.428	0.3636	35	0.757	0.2955
11	-0.002	0.4602	36	-0.247	0.5170
12	-0.642	0.6080	37	0.246	0.4034
13	-1.248	0.7330	38	0.021	0.4545
14	1.559	0.1648	39	0.348	0.3807
15	-0.986	0.6818	40	1.600	0.1591
16	0.402	0.3693	41	1.205	0.2159
17	-0.797	0.6420	42	-0.988	0.6818
18	0.324	0.3864	43	1.514	0.1705
19	0.375	0.3750	44	0.970	0.2557
20	-0.592	0.5966	45	1.068	0.2386
21	0.023	0.4545	46	1.278	0.2045
22	-0.002	0.4602	47	0.534	0.3409
23	-0.295	0.5284	48	-0.467	0.5682
24	-0.344	0.5398	49	-0.344	0.5398
25	-1.280	0.7386	50	0.349	0.3807

https://doi.org/10.46627/sipose.v5i1.330

According to Table 2, each element has indicators, which will be explained in Table 3 below about the tool grid for each element.

Table 3. The grid of entrepreneurship readiness instruments

Variable	Indicator	Sub Indicator	Question Number
Entrepreneurship Readiness	Entrepreneurial mental attitude	Forward oriented	3, 14, 34, 35
		Ability to formulate life/business goals	7, 9, 12, 43
		Initiative ability	2, 4
		Mental ability based on religion	1, 5, 6, 8
	Ability to take risks	Ability to form money capital	11, 42
		Ability to form capital goods	22, 47
	Creativity	Ability to innovate products	10, 26, 38
		Ability to innovate marketing	21, 44, 45, 46, 48, 49
	Motivation	Self-motivation ability	15, 16, 18, 24, 25
		Ability to motivate colleagues	36, 40, 41
	Entrepreneurial skills	Ability to build network	28, 30, 31
		Ability to influence people	20, 29
		Ability to collaborate	19, 37, 39
		Ability to communicate	17, 27, 31, 32, 50

Table 4 shows a rundown of parameter gauges and in general grain trouble from Table 2. The thing trouble parameter is the point on the capability scale where the likelihood of replying accurately is 0.5. These parameters are area parameters, which show the position of the grain characteristic bend around the capability scale. The more noteworthy the esteem of the thing trouble parameter, the more prominent the capacity required by the test taker to get a 50% chance of accurately replying the thing. The furthest left bend contains an area –2.075, and the furthest right bend has a zone of 0.159. Concurring to the classical test investigation, thing trouble ranges from 0.04 to 0.96.

Table 4. Summary of estimated parameters β and test item difficulties

	β	Item Difficulties (*p*)
Average	0.00006	0.465
Standard deviation	0.896	0.186
Minimum	-2.075	0.159
Maximum	0.159	0.858

Figure 2 shows the relationship between parameter β and item difficulty. The correlation between β and p value is 0.9989.

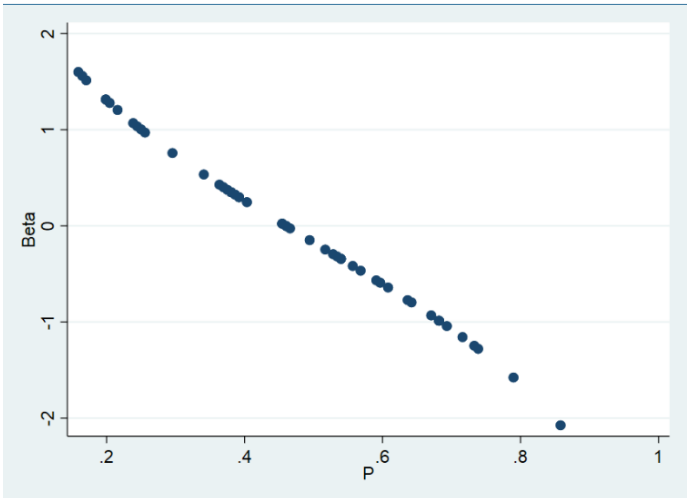


Figure 2. Curve characteristics of several test items

Discussion

An essential quality of a group of test items assessing a concept is that they are unidimensional. In Rasch analysis, if all items come together to create one scale, then the items are unidimensional. Adaptation of items was employed to assess unidimensionality (Keeves, 1999). Item fit was evaluated using the ConQuest program (Adam et al., 1996), which offers two fit measures: Infit Mean-Square (IMS) and Outfit Mean-Square (OMS). The findings of the analysis can be found in Table 5.

Table 5. Results of fit item analysis

No.	IMS	OMS	No.	IMS	OMS
1	1.00	0.95	26	1.00	0.97
2	0.99	1.01	27	1.00	1.00
3	0.90	0.92	28	0.93	0.90
4	0.98	0.98	29	0.99	0.98
5	0.98	0.91	30	0.99	0.90
6	0.93	1.35	31	0.97	1.09
7	0.95	0.86	32	0.96	0.89
8	0.93	1.11	33	0.94	1.06
9	0.95	0.94	34	0.97	0.94
10	1.03	0.98	35	1.01	0.95
11	0.97	0.92	36	0.98	0.93
12	1.17	1.00	37	1.13	1.00
13	1.13	1.14	38	1.06	1.13
14	0.79	1.04	39	0.88	1.00
15	0.96	0.87	40	0.98	0.93
16	0.99	1.15	41	1.00	1.04
17	0.95	1.01	42	0.97	1.00
18	0.91	1.19	43	0.92	1.07
19	1.00	1.24	44	0.99	1.12
20	1.07	0.79	45	1.05	0.85
21	0.96	1.00	46	0.96	1.01
22	1.00	1.01	47	1.00	1.00
23	1.00	1.09	48	1.00	1.07
24	0.98	1.25	49	0.98	1.21
25	0.96	1.14	50	0.99	1.12

When IMS and OMS values are close to 1.0, it indicates minimal distortion in the measurement system. A value of 1.0 is considered the norm. If the data aligns with the Rasch model, the anticipated value is 1.0. The IMS and OMS values in Table 4 demonstrate that all test items are practical for measurement. The Entrepreneurship Readiness Test assesses students' skills in innovative products and entrepreneurship within the IT professional sector in East Java. Thus, this test can serve as a guide for students aiming to graduate well-prepared for entrepreneurship. Rasch adjustment analysis is employed to examine if entrepreneurial readiness is unidimensional. Assessment based on equipment and facility statistics suggests that solid test items typically have root mean square values exceeding 0.5 but less than 1.5. A root mean square value surpassing 2.0 signifies more error than helpful information from a test item. The readiness test items displayed goodness-of-fit values ranging from 0.79 to 1.17 and retentions ranging from 0.79 to 1.35. The fit and hold index indicates the coherence between the data and the one-way Rasch model, suggesting that the entrepreneurship preparation test design is unidirectional.

CONCLUSION

The analysis conducted using Stata 16's Rasch program suggests that the test is sound overall. While some items may need adjustment or replacement, the majority are well-constructed. In

classical test theory, the item difficulty statistic is represented by the p-value, determined by the proportion of test takers who answered an item correctly. The parameter β indicates the position of the item characteristic curve on the ability scale and shows a high correlation of 0.9989 with p values. According to the Rasch model, all 50 entrepreneurial readiness assessments align with IMS values between 0.88 and 1.17 and OMS values between 0.86 and 1.35. The β parameter ranges from -2,075 to 0.159, while the difficulty parameter p ranges from 0.159 to 0.858. These findings suggest that attributes such as mindset towards entrepreneurship, risk-taking ability, creativity, motivation, and entrepreneurial skills play vital roles in assessing students' readiness for entrepreneurship.

RECOMMENDATION

The outcomes of the item analysis should be shared with item authors during the peer review phase, as doing so can enhance the caliber of forthcoming assessments.

ACKNOWLEDGMENT

The author expresses his gratitude to the academic community of Universitas Negeri Yogyakarta who have helped in financing as well as moral and material support so that this article can be made. We also do not forget to thank the faculty of Vocational, Universitas Negeri Yogyakarta and the head of the Department of Electrical & Electronics Engineering who have assisted in conducting research conducted in schools in East Java.

REFERENCES

Adams, R. J. & Khoo, S. (1996). *ACER Quest: The interactive test analysis system. Version 2.1.*

Allen, M. J. & Yen, W. M. (1979). *Introduction to measurement theory*. Brooks/Cole.

Andrich, D. (1998). *Rasch models for measurement*. Sage Publications.

Brooks, G. P., & Johanson, G. A. (2003). Test analysis program. *Applied Psychological Measurement*, *27*, 305-306.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists (multivariate applications series)*. Lawrence Erlbrum Associates.

Fischer, G. H., Fischer, & Molenaar, I. W. (1995). *Rasch models: Foundations, recent developments, and applications*. Springer-Verlag.

Hambleton, R. K. (2004). Theory, methods, and practices in testing for the 21st century. *Psicothema, 16*(4), 696–701.

Hambleton, R. K. (2000). Advances in performance assessment methodology. *Applied Psychological Measurement, 24*(4), 291–293.

Hambleton, R. K., Swaminatha, H., & Roger, H. J. (1993). Influence of item parameter estimation errors in test development. *Journal of Educational Measurement*, 2(30), 144-145.

Hulin, C. L., Dasgrow, F., & Parso, C. K. (1983). *Item response theory: Application to psychological measurement.* Dow Jones – Irwin.

Impara, J. C., & Plake, B. S. (1998). Teachers' ability to estimate item difficulty: A test of the assumptions in the angoff standard setting method. *Journal of Educational Measurement*, 35, 69-81.

Kaplan, R. M. & Saccuzo, D. P. (2009). *Psychological testing: Principles, applications and issue*. Brooks Cole Pub. Company.

Keeves, J. P., & Masters, G. N. (1999). Introduction. In G. N. Masters & J. P. Keeves, J. P. (Eds.), *Advences in measurement in educational research and assessment.* Pergamon an Imprint of Elsevier Science.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean ?. *Rasch Measurement Transactions, 16*(2), 878.

Mansyur, S., & Harun, R. (2015). *Asesmen pembelajaran di sekolah: Panduan bagi guru dan calon guru*. Pustaka Pelajar.

Mardapi, D. (2012). *Pengukuran penilaian dan evaluasi pendidikan*. Nuha Medika.

Novick, M. R. (1996). The axioms and principal results of clasical test theory. *Journal of Mathematical Psychology*, *3*, 1-18.

Russell, E. (2018). Making a math degree work for the workplace. *Higher Education Pedagogies, 3*(1), 403-416.

Smith, E. Jr. (2001). Evidance for the reliability of measures and the validity of measure interpretation: A rasch measurement perspective. *Journal of Applied Measurements, 2*(3), 281-311.

Subdirektorat Statistik Ketenagakerjaan. (2002). *Keadaan angkatan kerja di Indonesia*. Badan Pusat Statistik.

Suharno, Harjanto, B., Handayani, N. M., & Sutanti, N. (2018). Evaluation of five-day school program implementation using the model of contex, input, process, and product. *Jurnal Pendidikan Teknologi dan Kejuruan*, *24*(1), 155-161. http://dx.doi.org/10.21831/jptk.v24i1.18342

Suharno, Pambudi, N. A., & Harjanto, B. (2020). Vocational education in Indonesia: History, development, opportunities, and challenges. *Children and Youth Services Review*, 115, 1-8. https://doi.org/10.1016/j.childyouth.2020.105092

Suranto. (2006). Strategi pembelajaran dengan focused based education. *Jurnal Ilmiah Teknik Industri Universitas Muhammadiyah Solo, 4*(3).

United Nations Development Programme. (2018). *Human Development Report*.

Wu, M., Adams, R., & Wilson, M. R. (1998). *Conquest: Generalised item response modelling software*. Camberwell: Australian Council for Educational Research.

Author (s):

*Septian Rahman Hakim (Corresponding Author)
Department of Vocational Education, Postgraduate School,
Universitas Negeri Yogyakarta,
Jl. Karangmalang, Yogyakarta 55281, Indonesia
Email: septianrahmanhakim@uny.ac.id

Ekohariadi
Postgraduate School,
Universitas Negeri Surabaya,
Jl. Lidah Wetan, Surabaya 60213, Indonesia
Email: ekohariadi@unesa.ac.id

Suparji
Postgraduate School,
Universitas Negeri Surabaya,
Jl. Lidah Wetan, Surabaya 60213, Indonesia
Email: suparji@unesa.ac.id

Harmanto
Postgraduate School,
Universitas Negeri Surabaya,
Jl. Lidah Wetan, Surabaya 60213, Indonesia
Email: harmanto@unesa.ac.id

Elok Wiwin
Cabdin Pendidikan Wilayah Sidoarjo,
Jl. Jagir Sidoresmo V, Sidosermo, Kec. Wonocolo, Surabaya 60239, Indonesia
Email: buelok@gmail.com