JRERE 12 (1) 2023 67-75

Journal of Research and Educational Research Evaluation

http://journal.unnes.ac.id/sju/index.php/jere

The Development of Verbal Reasoning Assessment Instrument for Vocational School English Subject: Validity and Reliability

Ryzka Cahyaningrum[1]✉, Rudi Hartono[2], Ellianawati Ellianawati[3]

[1,2,3]Universitas Negeri Semarang, Indonesia

Article Info

History Articles
Received:
19 December 2023
Accepted:
17 January 2023
Published:
30 March 2023

Keywords:
Assessment instrument, reliability, validity, verbal reasoning

Abstract

Reasoning is one of the important skills that should be mastered especially for vocational school students. Vocational school students are prepared to be worked ready, so they should have a clear reasoning. Reasoning that uses language as the main tool is called verbal reasoning. This study aimed to analyze the validity and reliability of the development of verbal reasoning assessment instrument. Quantitative was used as the method of study. 5 validators and 249 eleventh-grade students were involved as the research subjects. A verbal reasoning assessment instrument was shared with the students for the data collection. The technique of analysis used was content validity using Aiken's V, construct validity using EFA and CFA, and reliability using KR-20. The results showed that content validity was validly stated by the mean of V value obtained was 0.853. EFA showed that the data was unidimensional. CFA indicated that the single-order model was supported by the data. Reliability indicated that the verbal reasoning assessment instrument was reliable by the r_i obtained was 0.881 for the small scale and 0.822 for the large scale. Therefore, the verbal reasoning assessment instrument developed was valid and reliable. The teachers can use verbal reasoning assessment instrument to measure or evaluate students' verbal reasoning in English subjects. Other than that, validity and reliability are the important parts in developing an instrument so this research may be used as a reference or information.

✉Correspondence Address :
Educational Research and Evaluation, Universitas Negeri Semarang, Indonesia 50229
E-mail : ryzkac@students.unnes.ac.id

p-ISSN 2252-6420
e-ISSN 2503-1732

INTRODUCTION

Reasoning that used language as a tool is called verbal reasoning (Wardita, 2017). Verbal reasoning is the process of using logic to a problem, information, or premise in the form of words that are carried out to a clear conclusion (Krumnack et al., 2011). Students can process, capture meaning, analyze, and conclude a reading or speech logically by honing verbal reasoning (Joseph et al., 2018). Verbal reasoning helps students to analyze the conditions faced from various sides, think clearly, make logical assumptions, and conclude information based on data or facts, so that students are not trapped in false information (Amir-Mofidi et al., 2012; Ramdhani, 2017). Therefore, verbal reasoning should be involved in the English learning process up to the assessment.

Assessment is a decision-making process related to students' learning outcomes which are used to investigate the students' knowledge and skills so that the teachers can make improvement and find out whether the material and the teaching methods provided are appropriate or not (Arikunto, 2018; Baird et al., 2017; Black & Wiliam, 2018). Doing test assessments produce long-term memories of subject matter more effectively than re-reading material learned (Butler & Roediger, 2007; Larsen et al., 2009, 2015; Pastötter & Bäuml, 2014; Roediger & Karpicke, 2006). Comprehension and verbal reasoning students related to verbal working memory (Barreyro et al., 2019).

Based on the observations at three Vocational Schools in Tangerang City, the teacher has never assessed verbal reasoning in English subjects due to the teacher did not have a verbal reasoning assessment instrument. Several verbal reasoning assessment instruments have been developed before but there have not been many updates regarding the use, content, form, and level of difficulty of verbal reasoning assessment instruments (Talman et al., 2020). Even in recent years, there has been a shift in content, where verbal reasoning assessment instruments not only contain verbal reasoning but also contain numerical and spatial assessments combined into aptitude tests or intelligence tests (Coalson et al., 2010; Roopesh, 2020; Trassi et al., 2019; Widhiarso, 2019). Furthermore, verbal reasoning assessment instruments are often used as recruitment tests or for medical purposes, whereas verbal reasoning assessments have the potential as formative and summative assessments in education (Basagni et al., 2017; Primrose et al., 2000; Widhiarso, 2019; Widhiarso & Haryanta, 2016).

Verbal reasoning has not been directly applied to English learning or validly and reliably applied to assessment (Bronkhorst et al., 2020). On the other hand, the ability to interpret and comprehend is one of the difficulties that students often experience in applying language skills and verbal reasoning. However, teachers can detect students' difficulties by conducting assessments, so that anticipation or improvement can be made (Seheri, 2016). Other researchers suggested conducting verbal reasoning assessments as well that can encourage students' understanding and verbal reasoning (Montejo & Jamon, 2018). Therefore, the verbal reasoning assessment instrument was developed.

The verbal reasoning assessment instrument developed must put attention to validity and reliability so that it can measure students' abilities properly and consistently. The development of the verbal reasoning assessment instrument referred to the material in Curriculum 2013 which was found to be related to verbal reasoning, namely analytical exposition text learned in eleventh grade. The analytical exposition text material was found to involve reasoning and critical thinking skills, so the use of verbal reasoning was assumed to be appropriate to the analytical exposition text (Kiptiyah, 2019; Rohayati, 2017; Umaroh et al., 2019). Analytical exposition text prioritized data or facts in the presented arguments.

Based on the explanation, this study aimed to prove and estimate the validity and reliability of the development of verbal reasoning assessment instruments. This study was expected to facilitate teachers in conducting verbal reasoning assessments and English learning evaluations. Therefore, the research question would be:

How did the content validity of the development of the verbal reasoning assessment instrument?

How did the construct validity of the development of the verbal reasoning assessment instrument?

How did the reliability of the development of the verbal reasoning assessment instrument?

METHODS

The quantitative method with the descriptive design was used in this study. The data were obtained first from the research subjects and analyzed to investigate the validity and reliability of the development of the verbal reasoning assessment instrument.

5 validators and 249 eleventh-grade students majoring in Office Governance Automation from four Vocational Schools in Tangerang City were involved as the research subjects. The students were divided into two scales, which were small-scale and large-scale. 74 students were included on a small scale, while 175 students were included on a large scale. The subjects were determined using purposive sampling and the Slovin formula.

The data were obtained through the expert validation sheets and the verbal reasoning assessment instruments. The expert validation sheet contained 20 statements with five Likert scales. The verbal reasoning assessment instrument contained 40 multiple-choice questions with five options. The score was dichotomous, which means that 1 for the correct answer and 0 for the wrong answer. The verbal reasoning assessment instrument consisted of five test components as the construct that would be proved.

The data were analyzed using content validity, construct validity, and reliability. Construct validity and reliability were used for the large scale as well. First, content validity used in this study was Aiken's V formula calculated using Microsoft Excel. The cut-off value was the V table, which was 0.80 for five validators and five scales.

Construct validity used in this study was factor analysis, which was Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA was used for the small-scale sample and CFA was used for the large-scale sample. EFA was analyzed using SPSS 29, while CFA was analyzed using Lisrel 8.80. The data used for construct analysis were the total score of each test component that was answered correctly.

EFA required a minimum value obtained to state that the assessment instrument was valid. The Kaiser-Meyer-Olkin (KMO) value should be > 0.50, Bartlett significance should be < 0.05, the eigenvalue showed > 1, scree plot and loading factor should be > 0.30 (Azwar, 2021; Retnawati, 2017). This study used Principal Component Analysis for the extraction.

CFA had cut-off values as well to prove the model. The proposed model was in a single order, which means the verbal reasoning assessment instrument was unidimensional. The cut-off values of CFA were Standardized Loading Factors (SLF) > 0.5, T-values > 1.96, and Goodness of Fit, that was Chi Square p-value > 0.05, GFI > 0.90, RMSEA < 0.08, SRMR < 0.08, AGFI > 0.90, TLI > 0.90, NFI > 0.90, CFI > 0.90, IFI > 0.90, and RFI > 0.95 (Cai et al., 2021; Diamantopoulos & Siguaw, 2000; Narimawati et al., 2020; Sun, 2005).

Last, the reliability used in this study was KR-20 because the data obtained was dichotomous. Reliability was estimated using SPSS 29.

RESULTS AND DISCUSSION

Content Validity

The validation sheet contained 20 statement that was assessed by five validators based on their expertise. This study involved material experts, media experts, and a practitioner as the validators. The result of content validity was presented in Table 1.

Table 1. Content Validity

Item	Vtable	Vcount	Description
1	0.80	0.80	Valid
2	0.80	0.80	Valid
3	0.80	0.80	Valid
4	0.80	0.80	Valid
5	0.80	0.95	Valid
6	0.80	0.85	Valid
7	0.80	0.90	Valid
8	0.80	0.90	Valid
9	0.80	0.95	Valid
10	0.80	0.80	Valid
11	0.80	0.85	Valid
12	0.80	0.85	Valid
13	0.80	0.90	Valid
14	0.80	0.85	Valid
15	0.80	0.80	Valid
16	0.80	0.90	Valid
17	0.80	0.85	Valid
18	0.80	0.80	Valid
19	0.80	0.80	Valid
20	0.80	0.90	Valid
Mean		0.853	Valid

Each item was stated as valid with all the Vcount and the mean of Vcount obtained was more than equal to Vtable. The mean obtained was 0.853 > 0.80. Therefore, the content validity was valid with the necessary revisions. The assessment instrument of verbal reasoning could be used for measuring verbal reasoning according to content validity.

Content validity plays an important role in the development of assessment instruments so that the concepts of the assessment instruments are following what you want to measure. Ad'hiya and Laksono used Aiken's V in their study as well related to analytical thinking skills with 20 statement items valid at 0.86 (Ad'hiya & Laksono, 2018). The reasoning is related to analytical thinking skills. The reasoning aspect was the outcomes of analytical thinking (Fadly, 2021). It showed that reasoning skills were an important skill that have to be mastered. One of the ways to achieve this is by practicing reasoning skills through verbal reasoning assessment.

Construct Validity

Construct validity was performed on a small scale and a large scale. The data used for EFA was small-scale data (74), while the data used for CFA was large-scale data (175).

Small Scale

The results of EFA were presented in Table 2., Table. 3, Figure 1., and Table 4. The results expected were the test components measure one attribute.

Table 2. KMO and Bartlett's Test

KMO	0.852
Bartlett's sig.	0.001

KMO obtained was 0.852 > 0.5 and Bartlett's significance obtained was 0.001 < 0.05. Therefore, the verbal reasoning assessment instrument was accepted to be analyzed. After KMO and Bartlett's significance were accepted, the eigenvalue was analyzed as could be seen in Table 3.

Table 3. Total Variance Explained

Initial Eigenvalues				Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.380	67.592	67.592	3.380	67.592	67.592
2	.554	11.090	78.682			
3	.438	8.757	87.439			

4	.355	7.101	94.540
5	.273	5.460	100.000

The highest eigenvalue obtained was 3.380 > 1 with 67% cumulative on component one. It meant the data measured one factor. Other components were less than 1, so it did not form the factor. Sometimes, the eigenvalue can create many components or factors but only measure one or two dominant factors. Therefore, it is recommended to analyze the scree plot. In this study, the scree plot was analyzed to ensure the data was unidimensional. The scree plot can be seen in Figure 1.

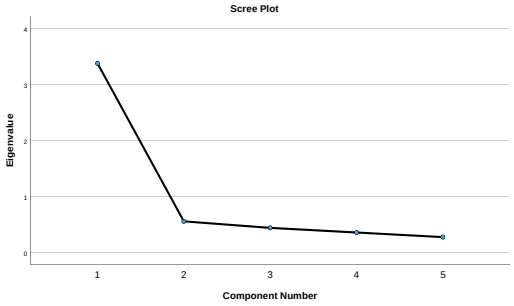

Figure 1. Scree Plot

The scree plot showed that the plot formed one line which meant the data was unidimensional. After the data was

found unidimensional, the data was extracted as can be seen in Table 4.

Table 4. Component Matrix

	Component 1
AN	.853
SC	.847
SG	.840
SY	.796
RC	.772

The loading factor of each test component was > 0.30. The highest loading factor was analogy (0.853) and the lowest loading factor was reading comprehension (0.772). Five test components were proved to measure one factor named verbal reasoning. Therefore, the verbal reasoning assessment instrument was found to measure one ability, which was verbal reasoning ability. EFA was used by Cavinez et al. as well to investigate Wechsler Intelligence Scale for Children-Fifth Edition in clinical cases (WISC-V) (Canivez et al., 2020). The results showed WISC-V was indicated to have a four-factor model. The results did not conform to the publisher. It showed that EFA could be used to validate existing instruments as this study did. This study collected the component test used through literature studies and validated it through EFA. Other researchers used EFA to validate existing instruments as well (Pattipeilohy & Widhiarso, 2018). The results showed that the instrument was multidimensional within two factors, namely verbal and nonverbal ability, with correlation and measuring general factor, namely cognitive reasoning skill. The verbal ability had the same component test with this study, such as synonym, analogy, and reading comprehension. The EFA used by other researchers included not only verbal skills but nonverbal skills as well such as mathematical or spatial so that the EFA results could be multidimensional. It is difficult to claim that the measured construct is purely unidimensional if a test is designed with a multidimensional content domain from the beginning (Azwar & Ridho, 2013). Meanwhile, this study only measured verbal reasoning, so the EFA results obtained could be unidimensional.

Large Scale

Construct validity was proved as well on a large scale to make sure that the unidimensional data was confirmed by CFA. The model supposed was a single factor. The results can be seen in Figure 2., Figure 3., and Table 5.

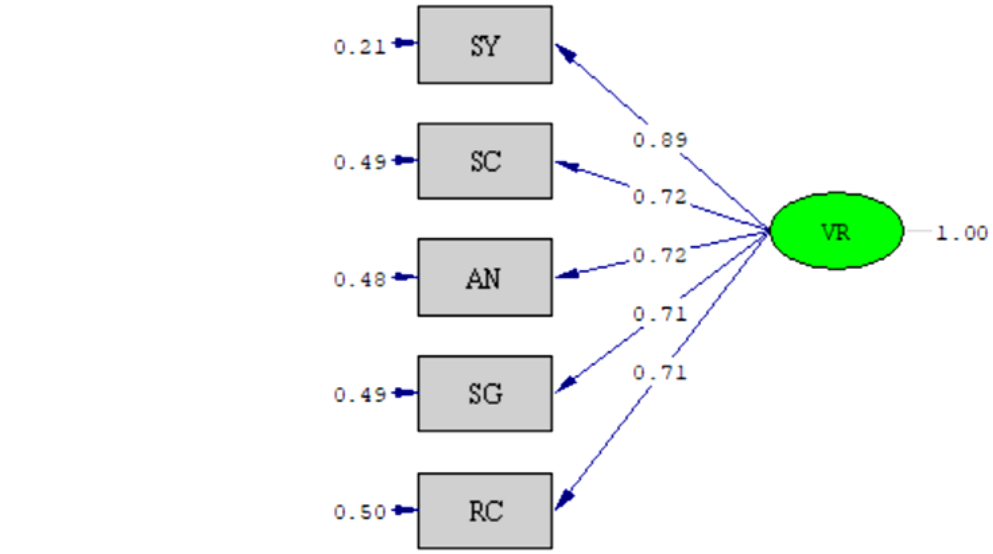

Figure 2. Path Diagram: Standardized Loading Factor

All SLF obtained was> 0.5. Synonym was 0.89, sentence completion was 0.72, analogy was 0.72, syllogism was 0.71, and reading comprehension was 0.71. After SLF was accepted, T-value was analyzed which can be seen in Figure 3.

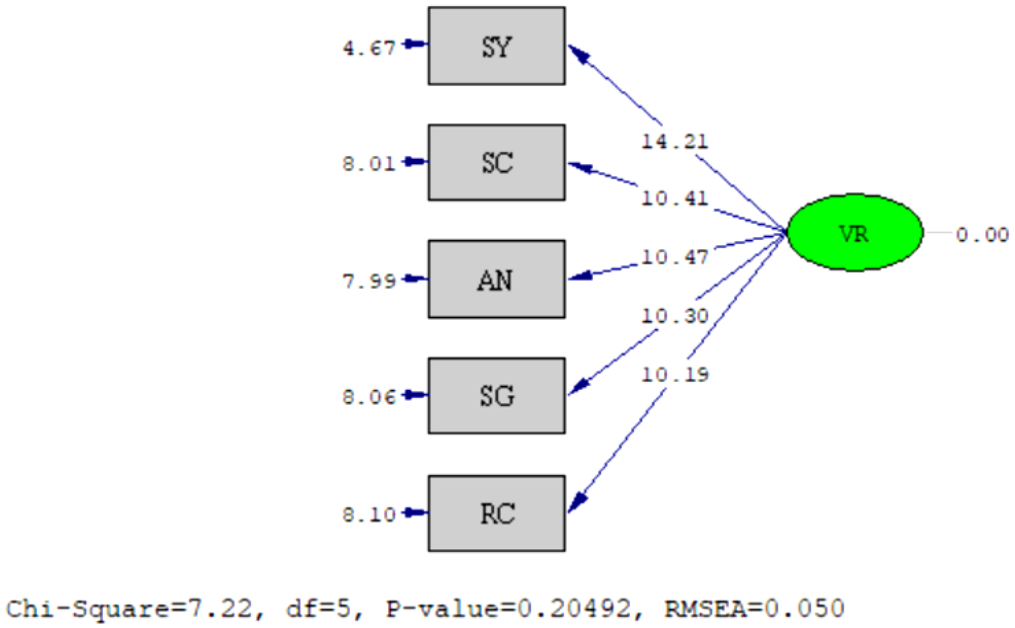

Figure 3. Path Diagram: T-values

All T-values obtained were > 1.96. Synonym T-value obtained was 14.21, sentence completion was 10.41, the analogy was 10.47, the syllogism 10.30, and the reading comprehension T-value obtained was 10.19. Therefore, SLF and T-values were supported by the data. The goodness of Fit was analyzed to investigate whether the model was fit or not. The results of Goodness of Fit can be seen in Table 5.

Table 5. Goodness of Fit

Criteria	Cut-off Value	Output
Chi-Square p-value	> 0.05	0.20
GFI	> 0.90	0.98
RMSEA	< 0.08	0.050
SRMR	< 0.08	0.025
AGFI	> 0.90	0.95
TLI	> 0.90	0.99
NFI	> 0.90	0.99
CFI	> 0.90	0.99
IFI	> 0.90	0.99
RFI	> 0.95	0.97

According to Goodness of Fit, all GOF criteria were accepted. Therefore, the single-order model was supported by the data. The assessment instrument of verbal reasoning was proved to measure one dimension, which was verbal reasoning. Sari et al. used CFA as well to analyze test instruments (Sari et al., 2022). They examined a test constructed by two dimensions, that were reasoning and proof, so the second order was used. The results revealed that the model was fit and supported by the data. Widhiarso used CFA to validate the structure of *Tes Potensi Akademik* (Widhiarso, 2019). The single-order and second-order were analyzed to prove which model can be fit. The results revealed that the test was unidimensional within a single order. The test measured one general factor or ability, which was reasoning skills. The result was the same as this study which

used single order and was supported by the data. The test component used was also the same as this study which used synonym, analogy, logical reasoning, and reading comprehension as the test component for measuring verbal reasoning. It can be concluded that the component test could be used to measure verbal reasoning theoretically.

Reliability

The reliability test was estimated using KR-20. The reliability test was estimated on small and large scales.

Small Scale

The result of reliability on a small scale can be seen in Table 6.

Table 6. Reliability Test on Small Scale

Reliability Test	Total Items
0.881	40

The reliability obtained on a small scale was 0.881 for 40 items.

Large Scale

Item analysis was conducted to the development of verbal reasoning assessment instrument that cannot discuss in this article. The item analysis caused an item to be taken out. Therefore, the reliability test for a large scale was estimated for 39 items. The results of reliability on a large scale can be seen in Table 7.

Table 7. Reliability Test on Large Scale

Reliability Test	Total Items
0.822	39

The reliability test gained for the large scale was 0.822.

This study was focused on validity and reliability. Differing from other studies, the verbal reasoning assessment instrument developed was only focused on the language matter and analytical exposition text matter. Therefore, it was expected to facilitate the teachers that want to measure students' verbal reasoning in English subjects.

CONCLUSION

Validity and reliability are important stages in assessment instrument development. It was conducted so that the assessment instrument used can measure validly and reliably. Validity and reliability employed can conform with the needs. This study used content and construct validity, while internal consistency reliability was employed.

The research question was about how the validity and reliability of the development of verbal reasoning assessment. The content

validity result stated that 20 statement of content validity was valid. The mean of all items of content validity was more than Vtable. It meant that the verbal reasoning assessment instrument was stated to be valid based on content validity.

Meanwhile, construct validity with EFA showed that the data formed a unidimensional on a small scale. Construct validity with CFA showed that the unidimensional model with single order was supported by the data on a large scale. Based on construct validity using factor analysis, the verbal reasoning assessment instrument was valid and measured one factor: verbal reasoning. Last, reliability results showed that the verbal reasoning assessment instrument was reliable. Therefore, the verbal reasoning assessment instrument developed can be used to measure verbal reasoning skills in eleventh-grade vocational school students. Item analysis should be discussed in another research. The practicality and effectivity of verbal reasoning assessment instrument should be conducted as well to complete the eligibility of verbal reasoning assessment instrument.

REFERENCES

Ad'hiya, E., & Laksono, E. W. (2018). Development and Validation of an Integrated Assessment Instrument to Assess Students' Analytical Thinking Skills in Chemical Literacy. *International Journal of Instruction*, *11*(4), 241–256.

Amir-Mofidi, S., Amiripour, P., & Bijan-Zadeh, M. H. (2012). Instruction of Mathematical Concepts Through Analogical Reasoning Skills. *Indian Journal of Science and Technology*, *5*(6), 2916–2922.

Arikunto, S. (2018). *Dasar-dasar Evaluasi Pendidikan*. Bumi Aksara.

Azwar, S. (2021). *Reliabilitas dan Validitas* (4th ed.). Pustaka Pelajar.

Azwar, S., & Ridho, A. (2013). Abilitas Komposit dalam Tes Potensi. *Jurnal Psikologi*, *40*(2), 127 – 142.

Baird, J. A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and Learning: Fields Apart? *Assessment in Education: Principles, Policy and Practice*, *24*(3), 317–350.

Barreyro, J. P., Injoque-Ricle, I., Formoso, J., & Burin, D. I. (2019). Computerized Working Memory Battery (BIMeT-V): Studying the Relation between Working Memory, Verbal Reasoning and Reading Comprehension. *Trends in Psychology*, *27*(1), 53–67.

Basagni, B., Luzzatti, C., Navarrete, E., Caputo, M., Scrocco, G., Damora, A., Giunchi, L., Gemignani, P., Caiazzo, A., Gambini, M. G., Avesani, R., Mancuso, M., Trojano, L., & De Tanti, A. (2017). VRT (Verbal Reasoning Test): a New Test for Assessment of Verbal Reasoning. Test Realization and Italian Normative Data from a Multicentric Study. *Verbal Reasoning Is a Complex, Multicomponent Function, Which Involves Activation of Functional ProceNeurological Sciences*, *38*(4), 643–650.

Black, P., & Wiliam, D. (2018). Classroom Assessment and Pedagogy. *Assessment in Education: Principles, Policy and Practice*, *25*(6), 1–25.

Bronkhorst, H., Roorda, G., Suhre, C., & Goedhart, M. (2020). Logical Reasoning in Formal and Everyday Reasoning Tasks. *International Journal of Science and Mathematics Education*, *18*(8), 1673–1694.

Butler, A. C., & Roediger, H. L. (2007). Testing Improves Long-Term Retention in a Simulated Classroom Setting. *European Journal of Cognitive Psychology*, *19*(4/5), 514–527.

Cai, L., Chung, S. W., & Lee, T. (2021). Incremental Model Fit Assessment in the Case of Categorical Data: Tucker–Lewis Index for Item Response Theory Modeling. *Prevention Science*, *2*.

Canivez, G. L., McGill, R. J., Dombrowski, S. C., Watkins, M. W., Pritchard, A. E., & Jacobson, L. A. (2020). Construct Validity of the WISC-V in Clinical Cases: Exploratory and Confirmatory Factor Analyses of the 10 Primary Subtests. *Assessment*, *27*(2), 274–296.

Coalson, D. L., Raiford, S. E., Saklofske, D. H., & Weiss, L. G. (2010). WAIS-IV. Advances in the Assessment of Intelligence. *WAIS-IV Clinical Use and Interpretation*, 3–23.

Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing Lisrel*. SAGE Publications.

Fadly, W. (2021). Profile of Students Analytical Thinking Skills in Learning Style for

Completing Substance Pressure Problems. *Jambura Physics Journal*, *3*(1), 1–15.

Joseph, E., Easvaradoss, V., Abraham, S., & Jain, S. (2018). Enhancing Verbal Reasoning of School Children through Chess Learning. *International Journal of Scientific Engineering and Science*, *2*(7), 1–3.

Kiptiyah, B. M. (2019). Critical and Creative Thinking in the Writing of the Exposition Text. *ISLLAC : Journal of Intensive Studies on Language, Literature, Art, and Culture*, *3*(1), 19–25.

Krumnack, A., Bucher, L., Nejasmic, J., Nebel, B., & Knauff, M. (2011). A Model for Relational Reasoning as Verbal Reasoning. *Cognitive Systems Research*, *12*(3–4), 377–392.

Larsen, D. P., Butler, A. C., Aung, W. Y., Corboy, J. R., Friedman, D. I., & Sperling, M. R. (2015). The Effects of Test-Enhaced Learning on Long-Term Retention in AAN Anual Meeting Courses. *American Academy of Neurology*, *84*(7), 748–754.

Larsen, D. P., Butler, A. C., & Roediger, H. L. (2009). Repeated Testing Improves Long-Term Retention Relative to Repeated Study: A Randomised Controlled Trial. *Medical Education*, *43*, 1174–1181.

Montejo, H. B., & Jamon, B. E. V. (2018). Verbal Comprehension and Verbal Reasoning of Graduated Senior High School Students: Critical Success Factors in College. *TESOL International Journal*, *13*(4), 55–66.

Narimawati, U., Sarwono, J., Affandi, A., & Priadana, S. (2020). *Ragam Analisis dalam Metode Penelitian*. ANDI.

Pastötter, B., & Bäuml, K. H. T. (2014). Retrieval Practice Enhances New learning: The Forward Effect of Testing. *Frontiers in Psychology*, *5*(APR), 1–5.

Pattipeilohy, F. W. C., & Widhiarso, W. (2018). *Pengujian Validitas Konstruk Tes Potensi Akademik Pascasarjana (PAPS) melalui Analisis Faktor Eksploratori*. *February*, 1–15.

Primrose, A. F., Fuller, M., & Littledyke, M. (2000). Verbal Reasoning Test Scores and Their Stability Over Time. *Educational Research*, *42*(2), 167–174.

Ramdhani, S. (2017). Analisis Kemampuan Penalaran Analogis Mahasiswa Pendidikan Matematika Dalam Persamaan Diferensial Ordo Satu. *Prisma*, *6*(2), 162–172.

Retnawati, H. (2017). *Validitas Reliabilitas dan Karakteristik Butir*. Parama Publishing.

Roediger, H. L., & Karpicke, J. D. (2006). *Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention*. *17*(3), 249–255.

Rohayati, D. (2017). Students' Critical Thinking in Writing an English Exposition Text. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, *82*.

Roopesh, B. N. (2020). Binet Kamat Test of Intelligence: Administration, Scoring and Interpretation -An In- Depth Appraisal. *Indian Journal of Mental Health*, *7*(3), 180.

Sari, Y. M., Retnawati, H., & Fiangga, S. (2022). The Construct Validity of Mathematical Reasoning and Proof Test Instrument Integrated with GeoGebra: Second-Order Confirmatory Factor Analysis. *Beta: Jurnal Tadris Matematika*, *15*(2), 104–118.

Seheri, N. H. (2016). *Language Reasoning and Verbal Reasoning Difficulties of Students in the four year BSc-Course at NWU* (Issue November). North-West University.

Sun, J. (2005). Assessing Goodness of Fit in Confirmatory Factor Analysis. *Measurement and Evaluation in Counseling and Development*, *37*(4), 240–256.

Talman, K., Vierula, J., Kanerva, A. M., Virkki, O., Koivisto, J. M., & Haavisto, E. (2020). Instruments for Assessing Reasoning Skills in Higher Education: A Scoping Review. *Assessment and Evaluation in Higher Education*, *46*(3), 376–392.

Trassi, A. P., de Oliveira, K. L., & Inácio, A. L. M. (2019). Reading Comprehension, Learning Strategies and Verbal Reasoning: Possible Relationships. *Psico-USF*, *24*(4), 615–624.

Umaroh, M., Akhyaruddin, & Setyonegoro, A. (2019). Penalaran Dalam Tulisan Ekspositif Siswa Kelas X SMK Negeri 5 Kabupaten Tebo. *Pena*, *6*(2), 1–20.

Wardita, R. (2017). The Effect of Instructional Approach and Verbal Reasoning on Students' Scientific Writing Ability. *Journal of Education, Teaching and Learning*, *2*(2), 121–127.

Widhiarso, W. (2019). Pembuktian Validitas terkait Struktur Tes Potensi Akademik Pascasarjana (PAPS) Universitas Gadjah Mada. *Jurnal Psikologi*, *46*(2), 145–162.

Widhiarso, W., & Haryanta. (2016). Comparing the Performance of Synonym and Antonym Tests in Measuring Verbal Abilities. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, *23*(3), 335–345.