Development and psychometric properties of the Faxtor Mechanical Aptitude Test (FMAT)

Fatiya Nurhadi Nasyiah[1,2*], *Afiya* Nur Nabila[1,3], *Nadim* Alfan Assa[1,3], *Naufal* Dzaki Hafizan[5], *Whisnu* Yudiana[3]

[1]Faxtor Indonesia, Indonesia
[2]Department of Psychology, Faculty of Education, Universitas Pendidikan Indonesia, Indonesia
[3]Department of Psychology, Faculty of Psychology, Universitas Padjadjaran, Indonesia.
[4]Center for Psychometrics Study, Faculty of Psychology, Universitas Padjadjaran, Indonesia.
[5]Engineering Physics Study Program, Faculty of Industrial Technology, Institut Teknologi Bandung, Indonesia

Abstract. The development of assessment tools in Indonesia that measure mechanical ability is severely limited. The assessment tool that was regularly administered to measure that ability was translated more than 40 years ago without any information regarding the standardization of the test. This study aimed to report on the development and evaluation of the initial psychometric properties of the Faxtor Mechanical Aptitude Test (FMAT). A total of 276 engineering students participated in the study, completing the 25-item FMAT along with measures of Fluid Reasoning (Gf) and Visual Processing (Gv). Analyses of reliability, internal structure validity via Confirmatory Factor Analysis (CFA), and convergent validity were conducted. Results indicated that the FMAT possesses good internal consistency reliability, its hypothesized four-domain structure was confirmed with an excellent model fit, and it demonstrated significant positive correlations with both Gf and Gv. These findings provide strong initial evidence supporting the FMAT as a reliable and valid instrument for assessing mechanical reasoning potential, with practical implications for talent selection and development in industrial settings.

1 Introduction

The capacity for mechanical reasoning is a fundamental competency that underpins range of critical performance indicators within industrial and technical work environments. The indicators include proficiency in technical diagnosis and problem-solving, the ability to work effectively without direct supervision, mastery of technical tools, comprehension of standard operating procedures, and the skill to modify instrumentation while adhering to safety standards [1,2]. Further reinforcing this relevance, research by Littleton demonstrated that individuals with high potential for mechanical reasoning tend to exhibit greater success in mechanical skills training [3]. In this context, the measurement of mechanical reasoning becomes essential, particularly for the purposes of selection, recruitment, competency development, and potential-based training in this field.

* Corresponding author: fatiya@faxtor.co.id

Mechanical aptitude is conceptualized as a process of mechanical inference, wherein an individual mentally deduces the consequences of interactions within a mechanical system based on available information [4]. This process is mediated by mental simulation, a cognitive mechanism that represents and manipulates the causal relationships between system components, both visually apparent (e.g., shape, position) and non-apparent (e.g., force, tension). This dual process indicates that mechanical reasoning is inherently reliant on more fundamental cognitive abilities, specifically logical reasoning and visual-spatial information processing. Referencing the Cattell-Horn-Carroll (CHC) theoretical framework, the aforementioned processes of inference and mental simulation correspond to two broad abilities: Fluid Reasoning (Gf) and Visual Processing (Gv). Gf underpins the ability to reason logically in novel situations, while Gv supports the ability to manipulate and interpret spatial information [5].

The measurement of mechanical aptitude itself has a long history dating back to the 1920s. During the early era, developed instruments included the Stenquist Mechanical Assembly Test and the Minnesota Mechanical Ability Test, among others. Over time, more modern tests were developed to achieve more reliable and valid assessments. Some of the popular and widely accepted instruments for measuring mechanical aptitude today include the Armed Services Vocational Aptitude Battery (ASVAB), the Wiesen Test of Mechanical Aptitude (WTMA), and the Bennett Mechanical Comprehension Test (BMCT). These tests are frequently used as industry standards or benchmarks in the development of new mechanical aptitude measurement instruments. Fundamentally, mechanical aptitude tests are developed based on basic physics concepts. Halliday et al. categorize physics into six major concepts: Mechanics, Matter and Heat, Oscillations & Waves, Electricity & Magnetism, Optics, and Modern Physics [6].

The availability of well-validated tools in the Indonesian setting is severely limited, despite the considerable industry need for mechanical aptitude testing. A number of studies in Indonesia aiming to measure mechanical ability often do not specify the instrument used or rely on measures from unpublished sources. Meanwhile, other studies use existing instrument but evaluation of their psychometric properties-particularly their validity and reliability-for the Indonesian population are often difficult to find or are not reported. There are also studies that measure mechanical aptitude indirectly through assessments of cognitive abilities considered relevant-such verbal reasoning, numerical ability, abstract reasoning, etc.- and the instrument is over 50 years old. This condition highlights a limitation and gap concerning the availability of standardized and valid mechanical aptitude tools for the Indonesian population.

In response to this gap, the Faxtor Mechanical Aptitude Test (FMAT) was developed. The core focus of the FMAT is to assess "aptitude"—an individual's potential to understand and learn mechanical principles. This allows organizations to identify high-potential talent by focusing on their capacity to learn rather than solely on pre- existing knowledge. Consequently, this research provides a significant benefit not only for industrial practitioners in selection and recruitment but also for practitioners in the education sector by offering the FMAT as a valid and reliable instrument. This study therefore seeks to document the development process and psychometric properties of the FMAT, focusing on its reliability, internal factor structure, and its correlation with other variables. The theoretical linkage to Gf and Gv provides the primary hypotheses for the convergent validity analysis in this study.

2 Methods

2.1 Participant

A cluster sampling method was used to collect 276 participants from five departments at a university in Bandung, Indonesia. The Civil Engineering department had 63 participants (23%), Electrical Engineering had 48 (17%), Automotive Engineering Education had 54 (20%), Mechanical Engineering Education had 53
(19%), and Building Engineering Education had 58 (21%). The sample comprised 163 males (59%) and 113 females (41%) with an age range of 18 to 22 years old (M:20.00, SD: 0.98). All participants were active students at the time of the study and participated voluntarily.

2.2 Instrument

Faxtor Mechanical Aptitude Test

The development of the FMAT followed several stages, including test conceptualization, test construction, test tryout, item analysis, and item revision [7].

These processes are detailed as follows:

Conceptualization. The conceptualization of the FMAT was conducted by an interdisciplinary team of researchers from the fields of Psychology, Psychometrics, and Engineering Physics. This process was grounded in a systematic literature review, wherein its developmental framework was synthesized from two primary sources: (1) fundamental concepts from classical physics, and (2) a content domain analysis of established mechanical assessment instruments (i.e., WTMA, BMCT, and ASVAB). A comparative mapping was conducted to identify the most representative and industrially relevant content domains. This process yielded four core domains: Mechanical Reasoning (MR), concerning problems related to reasoning about the solid matter in motion or at rest; Matter and Heat Reasoning (MH), concerning problems related to reasoning about liquids and related to temperature; Electricity Reasoning (ER), concerning problems related to reasoning about electricity; and Tool Knowledge (TK), concerning the understanding of measurement and general tools—which subsequently formed the basis for the FMAT's blueprint and item operationalization

Construction. The initial item writing process involved researchers with expertise in physics to ensure that each item was based on accurate and relevant physical principles. Items were focused on mechanical phenomena commonly found in everyday life to emphasize logical reasoning rather than prior knowledge. The draft items were then reviewed through an expert judgment process involving Subject Matter Experts (SMEs) from two fields: engineering experts to evaluate the physics-related content validity, and psychometric experts to evaluate the items' validity concerning the reasoning construct. Revisions were made in an iterative process based on expert feedback until the final initial item pool was established. This phase created an initial pool of 133 items.

Test Tryout and Item Analysis. A pilot study involving 86 engineering students from two universities in Bandung, Indonesia was conducted to empirically evaluate the psychometric quality of the developed items. The primary goal of this stage was to perform item reduction based on empirical data analysis. From an initial pool of 133 items, those exhibiting negative item discrimination or negative factor loadings were removed. After the analyses a total of 25 items were selected for inclusion in the final instrument.

Final Instrument. The final version of the FMAT consists of 25 multiple-choice items, each presenting four response options. The test score is a composite score, calculated by summing

the number of correct responses, which represents the respondent's overall mechanical reasoning ability.

Faxtor Cognitive Ability Test (FCAT)

To gather validity evidence based on relations to other variables, cognitive abilities were assessed using the Faxtor Cognitive Ability Test (FCAT). The FCAT is a computer-based assessment grounded in the Cattell- Horn-Carroll (CHC) theory. For this study, the test was used to measure two domains: Fluid Reasoning (Gf), defined as the ability to efficiently solve novel problems, and Visual Processing (Gv), defined as the ability to manipulate mental imagery. The Gf domain was measured using the 30-item Non-Verbal Reasoning (NVR) subtest ($\alpha = .85$), while the Gv domain was assessed using the 25-item Visualization (VS) subtest ($\alpha = .63$). The scores for each subtest were calculated by summing the correct responses and were subsequently converted into standardized scores for interpretation.

2.3 Data Analysis Technique

All data analyses were conducted using R version 4.3.2 to evaluate the psychometric properties of the FMAT. Internal consistency reliability was estimated using the psych package by calculating Cronbach's Alpha (α) and McDonald's Omega (ω). Internal structure validity was tested via Confirmatory Factor Analysis (CFA) using lavaan package, with model fit assessed using the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Standardized Root Mean Square Residual (SRMR), and the Root Mean Square Error of Approximation (RMSEA) indices, followed by an analysis of construct reliability (CR). Subsequently, convergent validity evidence was examined using psych package by compute Pearson's correlation coefficient (r) to analyze the relationships between FMAT scores and scores on Fluid Reasoning (Gf) and Visual Processing (Gv).

2.4 Procedure

Data for this study were collected in-person during group sessions guided by a test administrator. Prior to the session, each participant was explained the purpose of the study and was required to sign an informed consent form. After providing consent, participants completed a personal information form. The test administration was computer-based, with participants completing the battery of instruments in the following fixed order: the Faxtor Mechanical Aptitude Test (FMAT), followed by the NVR and VS subtests from FCAT. As an appreciation, each individual who completed the entire test battery received a reward in the form of an IDR 50,000 e-wallet balance.

3 Result

3.1 Descriptive Statistic

Descriptive statistics for all measured variables are presented in Table 1. Skewness and kurtosis values indicated that all variables were approximately normally distributed. This conclusion is based on the values falling within the recommended range of ±1.5 for skewness and ±3.0 for kurtosis [8].

Table 1. Descriptive Statistics and Inter-Variable Correlation Matrix

Variable	N Item	M	SD	Range Score	Skew	Kurt	1	2	
1. Mechanical Aptitude (FMAT)	25	14.0	4.24	3-23	-0.36	-0.39	-		
2. Fluid Reasoning (NVR)	30	95.6	23.7	55-136	-0.94	0.35	0.355***	-	
3. Visual Processing (VS)	25	101	16.9	55-148	-0.056	-0.04	0.407***	0.438***	-

*** *p<.001*

3.2 Internal Consistency

Internal consistency reliability analysis was conducted on the FMAT total scores to evaluate the extent to which the test items consistently measure the same underlying construct. The results yielded a Cronbach's Alpha (α) coefficient of .74 and a McDonald's Omega (ω) coefficient of .75. As both coefficients exceed the recommended .70 threshold for acceptable reliability [9], these findings suggest that the FMAT possesses good internal consistency.

3.3 Validity Evidance Based on The Interbal Structure

A first-order Confirmatory Factor Analysis (CFA) was conducted to test the internal structure validity of the FMAT. This analysis aimed to confirm whether the four domains (Mechanical Reasoning, Matter and Heat Reasoning, Electricity Reasoning, and Tool Knowledge) significantly represented a single latent construct of Mechanical Aptitude, as visualized in Figure 1. The results indicated an excellent model fit to the data demonstrated by a non-significant Chi-square test ($\chi2(2) = 2.80$, $p > .05$) and supported by other strong fit indices (CFI = 0.995, TLI = 0.986, SRMR = 0.0175, and RMSEA = 0.0381). All of these indices met the criteria for an ideal goodness-of-fit [10], suggesting that the empirical data presented supports the hypothesized four-domain model. Furthermore, the standardized factor loading for each domain were as follows: Mechanical Reasoning (λ = .80), Matter and Heat Reasoning (λ = .64), Tool Knowledge (λ = .49), and Electrical Reasoning (λ = .41). All loadings in this range are considered significant, as they exceed the minimum threshold of .35 recommended for a sample size of over 250 [11]. The construct reliability (CR) for this model was .70, which meets the recommended standard [11], confirming that the four domains consistently measure the same latent construct.

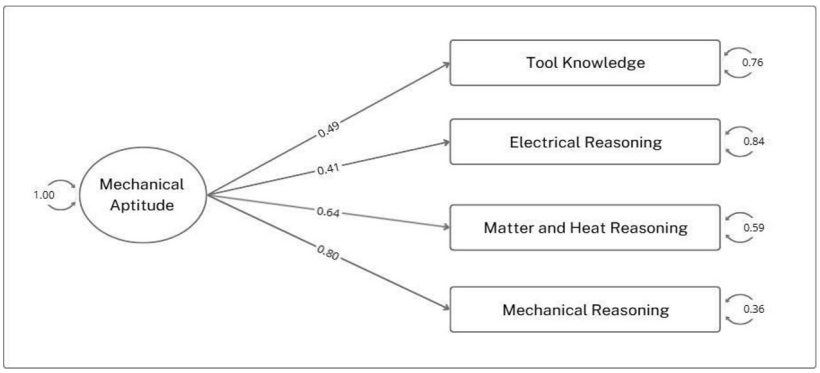

Fig 1. Confirmatory Factor Analysis Model

3.4 Validity Evidence Based on Relation to Other Variables

Pearson product-moment correlations were computed to assess the correlation between Mechanical Aptitude and related cognitive abilities, namely Fluid Reasoning (Gf) and Visual Processing (Gv). As presented in Table 1, FMAT scores demonstrated significant, positive correlations with both Fluid Reasoning (Gf; $r = .36$, $p < .001$) and Visual Processing (Gv; $r = .41$, $p < .001$). These findings are consistent with theoretical expectations and provide strong evidence for the construct validity of the FMAT. Notably, the magnitude of the correlation with Gv was slightly higher, aligning with the central role of Visual Processing in mechanical tasks. The results indicate that individuals with stronger logical reasoning (Fluid Reasoning) and a greater ability to mentally manipulate visual information (Visual Processing) tend to perform better on the mechanical aptitude test.

4 Discussion

This study aimed to present initial psychometric evidence for the Faxtor Mechanical Aptitude Test (FMAT), a new instrument designed to measure mechanical reasoning potential. Overall, the findings indicate that the FMAT is a promising measure with good psychometric properties. The initial findings revealed that the FMAT possesses good internal consistency reliability, suggesting that its items are consistent and can be depended upon to produce stable and reliable scores for measuring mechanical aptitude [9].

Validity evidence based on the internal structure using Confirmatory Factor Analysis (CFA) provides strong support for the FMAT's theoretical framework. The results, which indicated an excellent model fit, confirm that the hypothesized four-domain structure (Mechanical Reasoning, Matter and Heat Reasoning, Electricity Reasoning, and Tool Knowledge) is empirically proven to represent a single latent construct of Mechanical Aptitude. Unlike previous psychometric studies that often only report on overall model fit [12], this study also presents the factor loadings for each domain. This finding is crucial as it demonstrates that the FMAT is a coherently structured instrument aligned with its theoretical design.

An examination of the factor loadings reveals that Mechanical Reasoning emerged as the domain with the highest factor loading, suggesting it is the most potent indicator of the general mechanical aptitude construct. Conversely, Electrical Reasoning exhibited the lowest factor loading. A possible explanation for this discrepancy may lie in the observability of the phenomena being measured. The Mechanical Reasoning domain involves systems whose objects of reasoning are directly observable, as it is closely linked to everyday activities or hands-on experience such as repairing equipment or common carpentry tasks [13]. In contrast, the working system in electrical systems is more abstract and not directly observable [14], potentially engaging a slightly different reasoning process.

The final findings indicate that the FMAT showed positive and significant correlations with both Fluid Reasoning (Gf) and Visual Processing (Gv). These results provide strong validity evidence based on relations to other variables. The relationship with Gf is conceptually grounded in the nature of mechanical aptitude itself. Solving mechanical problems requires logical thinking and flexible problem-solving, a process described as inferring the behavior of a mechanical system [4]. These abilities are core functions of Fluid Reasoning as defined within the CHC theory of cognitive abilities [5]. The results of this study empirically support this link, aligning with recent research that found that fluid intelligence is strongly associated with a range of technical skills [15]. Additionally, the significant correlation with Visual Processing (Gv) further strengthens the validity of the FMAT. This result is consistent with the theoretical view that mechanical reasoning is not solely a logical process but also inherently involves the mental manipulation and simulation

of objects [4]. This finding aligns with earlier research indicating a relationship between visualization ability and success in solving mechanical problems [4]. Taken together, this pattern of results confirms the research hypothesis and substantially supports the construct validity of the FMAT as a measure of mechanical aptitude.

5 Conclusion

This study concludes that the Faxtor Mechanical Aptitude Test (FMAT) is an instrument with sufficient initial evidence of reliability and validity for measuring mechanical reasoning potential. A synthesis of the findings demonstrates that the FMAT is internally consistent, possesses a coherent four-domain structure aligned with its theoretical design, and is significantly correlated with core cognitive abilities such as fluid reasoning and visual processing. Therefore, the development of the FMAT provides a significant contribution by offering a contextually relevant tool with sufficient initial evidence of its psychometric properties to address the gap in standardized mechanical aptitude assessment in Indonesia.

6 Implication

The findings of this study offer several practical implications, particularly for practitioners in human resources. In the context of selection and recruitment, the FMAT can be used as an objective tool to identify candidates with a high potential for success in roles demanding mechanical reasoning. It allows organizations to assess an applicant's aptitude for learning and adapting to technical challenges, rather than relying solely on prior knowledge. Additionally, the FMAT can be used for employee development by identifying an individual's training needs.

Beyond its practical applications, the findings also suggest directions for future research. First, a predictive validity study represents a crucial step to determine whether FMAT scores can significantly predict important criteria, such as success in technical education or real-world performance ratings. Second, a current limitation is that the FMAT's output is confined to a composite score representing general mechanical aptitude. Future research could focus on developing the FMAT to provide a more comprehensive profile capable of differentiating abilities across its specific domains. Finally, as the current data is limited to engineering students, future validation studies should also include more diverse samples, such as technicians from various industrial sectors or non-engineering students to enhance the generalizability of the findings.

Acknowledgements: The Research Team would like to thank all those who have supported and contributed to the completion of this research, especially the item writer team and all students who have participated in this research.

Funding: This research was funded and supported by Faxtor Indonesia

Data Availability Statement: Anonymized data underlying the findings can be provided on reasonable request, conditional on a data-use agreement and relevant approvals.

Authors' Contributions Statement: FNN performing research, data curation, analyzing data and writing the paper; ANN performing research, data curation; NAA designing the study, review & editing the manuscript; NDH performing research; WY review & editing the manuscript. All authors read and approved the final version of the manuscript.

Competing Interest: FNN, ANN, and NAA are employees of Faxtor Indonesia; NDH collaborated externally and has no relationship with Faxtor; WY supervised the study. The authors declare no additional competing interests.

References

1. R. Kolz, L. A. Mcfarland, and S. B. Silverman, Cognitive Ability and Job Experience as Predictors of Work Performance, J. Psychol. **132**, 539 (1998) https://doi.org/10.1080/00223989809599286
2. P. M. Muchinsky, Validation of intelligence and mechanical aptitude tests in selecting employees for manufacturing jobs, J. Bus. Psychol. **7**, 373 (1993)
3. I. T. Littleton, Prediction in auto trade courses, J. Appl. Psychol. **36**, 15 (1952) https://doi.org/10.1037/h0056684
4. M. Hegarty, Mechanical reasoning by mental simulation, Trends Cogn. Sci. **8**, 280 (2004) https://doi.org/10.1016/j.tics.2004.04.001
5. B. Schneider, M. G. Ehrhart, and W. H. Macey, Organizational Climate and Culture, Annu. Rev. Psychol. **64**, 361 (2013) https://doi.org/10.1146/annurev-psych-113011-143809
6. R. Resnick, D. Halliday, and J. Walker, *Fundamentals of Physics*, 10th ed. (Wiley, 2014)
7. R. J. Cohen, W. J. Schneider, and R. Tobin, *Psychological Testing and Assessment: 2025 Release ISE* (McGraw Hill, Place of Publication Not Identified, 2025)
8. B. Tabachnick and L. Fidell, *Using Multivariate Statistics* (Pearson, NY, NY, 2019)
9. R. Kaplan and D. Saccuzzo, *Psychological Testing: Principles, Applications, and Issues* (Wadsworth Publishing, Boston, MA, 2018)
10. D. Hooper, J. Coughlan, and M. R. Mullen, Structural Equation Modelling: Guidelines for Determining Model Fit, Electron. J. Bus. Res. Methods **6**, pp53-60 (2008)
11. J. F. H. Hair Jr, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis* (Pearson, Upper Saddle River, NJ, 2010)
12. D. Bairaktarova and D. Reeping, Development and Psychometrics of a Freely Available Mechanical Aptitude Test, Int. J. Eng. Educ. **35**, 1839 (2019)
13. A. L. Pereira and M. H. Miller, Gender comparisons of mechanical aptitude, prior experiences, and enginering atitudes for mechanical enginering student, J. Women Minor. Sci. Eng. **18**, (2012) https://doi.org/10.1615/JWomenMinorScienEng.2013003830
14. K. S. Taber, T. de Trafford, and T. Quail, Conceptual resources for constructing the concepts of electricity: the role of models, analogies and imagination, Phys. Educ. **41**, 155 (2006) https://doi.org/10.1088/0031-9120/41/2/006
15. Y. Ganzach, Antecedents of Interest and the Investment of Fluid Intelligence in the Formation of Crystalized Intelligence, Front. Psychol. **12**, (2021) https://doi.org/10.3389/fpsyg.2021.679504