Scale Construction of The Discipline Scale For Indonesian Air Force (*TNI Angkatan Udara*) Military Personnel (IMPI-D32)

Yashinta Levy Septiarly[1], Saifuddin Azwar[2]

Study Program of Psychology, Universitas Dhyana Pura, Indonesia[1]
Faculty of Psychology, Universitas Gadjah Mada, Indonesia[2]

yashinta.levy@undhirabali.ac.id

Abstract

This study aims to create a discipline measuring tool with good psychometric properties for Indonesian Air Force Military Personnel (IMPI-D32). It was conducted using an integration of self-discipline, motivation to obey, and discipline theories in a military context. The IMPI-D32 consists of 32 items and its construction are based on the Classical Theory Test. The item's content validity was estimated using the Aiken V index at a minimum of .690. Furthermore, the item discriminations showed that the corrected item-total correlation for all items is above .3. Meanwhile, the reliability estimation with the Spearman-Brown formula indicated a satisfactory coefficient value of .916. The standard error of measuring the IMPI-D32 was 3.439 with the confidence of true score interval between X-6 to X+6.

Keywords: discipline, Indonesian air force, scale construction, psychometric

Abstract

Penelitian ini bertujuan untuk membuat alat ukur kedisiplinan dengan sifat psikometrik yang baik bagi Personil Militer Angkatan Udara Indonesia (IMPI-D32). Hal tersebut dilakukan dengan menggunakan integrasi disiplin diri, motivasi untuk taat, dan teori disiplin dalam konteks militer. Skala IMPI-D32 terdiri dari 32 item, dan konstruksinya didasarkan pada Tes Teori Klasik. Validitas isi item diperkirakan menggunakan indeks Aiken V minimal .690. Pada item diskriminasi menunjukkan bahwa korelasi item-total terkoreksi berada di atas .3 pada setiap item. Sedangkan estimasi reliabilitas dengan rumus Spearman-Brown menunjukkan nilai koefisien yang memuaskan sebesar .916. Standar error pengukuran IMPI-D32 adalah 3.439 dengan kepercayaan interval skor benar antara X-6 sampai X+6.

Kata kunci: angkatan udara Indonesia, disiplin, konstruksi skala, psikometris

Introduction

The Air Force of Tentara Nasional Indonesia (TNI), as a means of national defense responsible for maintaining the nation's sovereignty, territorial integrity, and safety, requires special characteristics to carry out its functions. According to Undang-Undang Number 34, Article 22 of 2004, and Number 25 of 2014 concerning the Indonesian National Army and the Military Discipline Law, soldiers require a high level of discipline to carry out their duties and functions effectively.

Preliminary studies have been conducted to strengthen the evidence that discipline is an important characteristic of soldiers. In this study, 144 Air Force personnel were asked to write down and describe the important characteristics of soldiers. The results showed that discipline is the most important characteristic (7.36%), followed by loyalty (15.14%), perseverance (13.89%), and having a leadership spirit (10.00%). Therefore, it is essential for Air Force TNI personnel to possess these attributes to develop the Indonesian Military Personality Inventory (IMPI) properly.

IMPI is a comprehensive measurement instrument divided into four series, namely IMPI-L, IMPI-P32, IMPI-D32, and IMPI-K, each of which measures the attributes of loyalty, persistency, discipline, and leadership. This study constructed a discipline scale for the Indonesian Military Personality Inventory-Discipline32 (IMPI-D32). Other IMPI measuring devices were extracted in a separate study.

The discipline theory used in IMPI-D32 integrates the concept of self-discipline (Duckworth & Seligman, 2006; Taylor, Kuo, & Sullivan, 2002), motivation to obey (Spector, 1978), and its concept in a military context. Self-discipline theory and motivation to obey are chosen because both concepts can explain the discipline described by soldiers in preliminary studies. This study defines discipline as a character that enables individuals to concentrate on achieving goals in various situations. It suppresses or inhibits one's ability to quickly comply with rules and ethics, overcome impatience for immediate gratification, and motivate obedience to rules and authority, which comes from within an individual as an intrinsic control. Several military personality assessments have also been developed and used in various countries. The ANM Toolkit consists of three measuring tools: a personality test, the need for external stimulation, and motivation used in Romania (Mihai, Corneliu & Sandina, 2014). Others include the Test of Personal Intelligence (TOPI) used to determine problem-solving areas based on the big five personality theory and used in the United States (Mayer, Panter, & Caruso, 2017), as well as the 5PF mile 2.0 and Rorschach, which uses two personality scales with norms made from soldiers in Norway (Hartmann, Sunde, Kristensen, & Martinussen, 2003).

Methods

This scale is constructed by determining aspects and behavioral indicators, which are arranged into a blueprint, to evaluate the item and response format. The formulation consists of four aspects, and each consists of two behavioral indicators represented by four items. Therefore, the IMPI-D32 measuring instrument consists of 32 items with its blueprint shown in **Table 1**.

A total of 72 items were created in a bid to have satisfactory psychometric qualities and used to determine the readability testing, estimate content validity, and field tests. This is in accordance with Azwar's (2017) research stating that to get the desired number of items and have good psychometric quality, it is essential to have 2 to 3 times the number of items planned in the blueprint.

Subsequently, the psychometric quality of the IMPI-D32 scale was conducted with a readability test to determine whether the sentence or terms used are easily understood in the military environment. The data source was the TNI Air Force personnel in Lanud Ngurah Rai (military air force base), which consisted of 30 people.

Content validity can be used to test the validity of a measurement instrument (Thorndike, 1997). The content validity estimation was carried out using Aiken's V index. Data were collected from 25 Masters in Psychology students of Universitas Gadjah Mada. The use of students as raters was conducted in accordance

with Aiken's (1985) research stating that two independent conditions can be used to measure instrument studies. Aiken's V, used as a limit for receiving items with 25 raters, is .67.

Table 1. The Blueprint of IMPI-D32

Aspect	Behavioral Indicator	Number of Items Desired	Number of Items Created
Able to concentrate on achieving goals in various situations	Having the ability to ignore outside interference to stay focused on achieving organizational goals	4	9
	Having the ability to ignore internal interference to stay focused on achieving organizational goals	4	9
Having the ability to suppress or hinder hasty responses to comply with rules and ethics	Able to consider alternative responses to the situation at hand, to avoid endangering themselves, especially the organization	4	9
	Behaving according to the rules and moral standards	4	9
Having the ability to overcome impatience	Not easily satisfied without achieving organizational goals	4	9
	Willing to overcome impatience to achieve long-term goals	4	9
Have the inner motivation to obey rules	Inability to make the punishment as the primary motivation to be obedient	4	9
	Have a desire to obey the rules and authority without coercion from outside themselves	4	9
	Total	32	72

Sources: Personal data

The item discrimination is estimated using the corrected item-total correlation coefficient ($r_{i(x-i)}$) with an acceptance criteria of $r_{i(x-i)} \geq .30$. Response data were obtained from 229 TNI soldiers of the Bintara Air Force who were learning at the Sekolah Pembentukan Perwira (Setukpa) Adi Soemarmo Air Force Base. The data collected were tested to determine the reliability and validity of the measurement results using the Spearman-Brown formula to produce a more stable estimate.

The Spearman-Brown formula can be used when parallel test assumptions can be proven. The standard error of measurement (S_e) and true score is estimated using a factor analysis technique during this stage. This is carried out to prove that the IMPI-D32 scale has one dominant factor. The estimated reliability and construct validity of the measurement results was conducted on 32 items of IMPI-D32. Factor analysis techniques are used to provide empirical evidence of testing dimensionality measurements.

Results and Discussion

Readability Test

The readability test results showed that participants generally stated that the sentences used in measuring instruments are easily understood. However, some suggested the need for changes in accordance with the context of the Indonesian Air Force, which led to the revision of 9 items.

Item Content Validity Estimation

Seventy-two items had the lowest and highest Aiken V index values of .690 and .990, respectively. The average index of Aiken's V measuring instrument is .849, and based on the receipt points, all items of this scale met the content validity criteria. Therefore, the IMPI-D32 discipline scale items have good relevance between indicators and items.

Item Discrimination Estimation

Based on the limitation, a total of 42 items that meet the criteria of $r_{i(x-i)} \geq .30$ in each indicator were received, while 30 were rejected. Each indicator is chosen four items with the highest $r_{i(x-i)}$, therefore 32

items were obtained according to the specified blueprint. **Table 2** shows the selected items marked with the numbers in bold.

Table 2. Items with Discrimination Index above .30

Indicator Number	Indicator	Number	$r_{i(x-i)}$
Indicator 1	Having the ability to ignore outside interference to stay focused on achieving organizational goals	**2U**	.445
		8U	.432
		1F	.353
		9U	.352
		5U	.322
		3F	.309
Indicator 2	Having the ability to ignore internal interference to stay focused on achieving organizational goals	**10U**	.493
		13F	.375
		16U	.349
		12F	.338
		14F	.325
Indicator 3	Able to consider alternative responses to the situation at hand so as not to potentially endanger themselves, especially the organization	**22F**	.394
		27U	.377
		25F	.375
		20F	.350
		19F	.341
Indicator 4	Behave according to the rules and moral standards	**34U**	.549
		33U	.543
		35U	.432
		31U	.429
		29F	.384
		32U	.355
		28F	.342
Indicator 5	Not easy to feel satisfied if organizational goals have not been achieved	**43F**	.409
		41F	.386
		37U	.357
		45U	.329
Indicator 6	Willing to overcome impatience to achieve long-term goals	**54F**	.521
		52F	.478
		53F	.442
		50U	.404
		47U	.381
		49U	.366
Indicator 7	Not make the punishment as the main motivation to be obedient.	**61U**	.482
		59F	.401
		58U	.314
		63F	.314
Indicator 8	Have a desire to obey the rules and authority without coercion from outside themselves	**71U**	.527
		65F	.391
		67U	.387
		69U	.318

Where:
1. The items in bold are the items selected in the IMPI-D32
2. F = favorable; U = unfavorable

Scale Format and Manual

Table 3. shows the format and final numbering of measuring instruments, where 32 selected items were selected and re-numbered.

Table 3. Change in Item Number of IMPI-D32 Scale

Indicator	Item Number		Indicator	Item Number	
	Initial number	New number		Initial number	New number
Indicator 1	1F	1F	Indicator 5	37U	17U
	2U	2U		41F	18F
	8U	3U		43F	19F
	9U	4U		45U	20U
Indicator 2	10U	5U	Indicator 6	50U	21U
	12F	6F		52F	22F
	13F	7F		53F	23F
	16U	8U		54F	24F
Indicator 3	20F	9F	Indicator 7	58U	25U
	22F	10F		59F	26F
	25F	11F		61U	27U
	27U	12U		63F	28F
Indicator 4	31U	13U	Indicator 8	65F	29F
	33U	14U		67U	30U
	34U	15U		69U	31U
	35U	16U		71U	32U

The final format of the scale is made as part of the measurement tool construction, processed, and ready for use. This scale is titled Indonesian Military Personality Inventory-Discipline32 (IMPI-D32) because it is used to measure the discipline of Indonesian Air Force soldiers in 32 items. Furthermore, manuals were designed to guide users' administration and interpretation of scores. The IMPI-D32 scale manual contains information on introductions, basic constructs, users, targets, grain format, administration, work instructions, scoring, interpretations, and attachments.

Reliability and Standard Error Estimation

The reliability estimation of the IMPI-D32 was conducted using the Spearman-Brown formula and divided into two parts to determine the distribution of aspects and indicators in each section. The division of the IMPI-D32 into two parts is shown in **Table 4**.

Table 4. The Division of The IMPI-D32

Indicator	Section	
	One (y1)	Two (y2)
Indicator 1	1	2
	3	4
Indicator 2	5	7
	6	8
Indicator3	9	11
	10	12
Indicator 4	14	13
	16	15
Indicator 5	19	17
	20	18
Indicator 6	22	21
	23	24
Indicator 7	25	26
	27	28
Indicator 8	29	31
	30	32
Discrimination Index	6.567	6.479
Mean	62.83	61.97
Variance	33.39	32.64
Correlation	.845	

Table 4 shows that the means, variance and correlation of the two sections are equivalent with significant correlation. Therefore, it can be concluded that parallel assumptions are determined using the reliability estimation and Spearman-Brown formula. The following are the results of the reliability estimation using the Spearman-Brown formula:

S-B = $r_{xx'}$ = 2 (r_{y1y2}) / (1+ r_{y1y2})
S-B = $r_{xx'}$ = 2 (.845) / (1+.845)
S-B = $r_{xx'}$ = 1.690 / 1.845
S-B = $r_{xx'}$ = .916
where : $r_{xx'}$ = Estimated reliability with the Spearman-Brown formula
          r_{y1y2} = Correlation of section one and two

At this stage, the standard error of measurement (S_e) was calculated to determine the estimated reliability value of Spearman-Brown and standard deviation (S_x) values of .916 and 11.037, as follows:

$S_e = S_x\sqrt{(1 - r_{xx'})}$
$S_e = 11.037\sqrt{(1 - .916)}$
$S_e = 11.037 \sqrt{.084}$
$S_e = 11.037 (.290)$
$S_e = 3.199$

Using S_e = 3.199, a true confidence score interval (Γ) can be estimated to determine the true scores free from measurement errors (Azwar, 2017). True scores can be estimated using visible scores obtained from measurements with the following formula:

$\Gamma = X \pm z_{\propto/2} (S_e)$
where: $z_{\propto/2}$ = the magnitude of z at the significance level $\propto/2$

At a 95% confidence level (.95), $\propto$ = 1 - 0.95 = .05, therefore $\propto/2$ = .025. Therefore, with reference to the normal deviation table, the value of z =

±1.96 at the right and left ends of the normal distribution. Therefore, the true confidence score interval is estimated as follows:

$\Gamma = X \pm 1.96 (3.199)$
$\Gamma = X \pm 6.270$

The true score estimation results indicate that the pure score (Γ) for the X value obtained from the measurement has a 95% chance of being within the range -6.270 and + 6.270. This research was carried out using a non-decimal score with an estimated pure score in the range X ± 6.

Construct Validity Estimation

The factor analysis performed on 32 IMPI-D32 items indicate that the scale is dominant or a single dimension. The Kaiser-Meyer-Olkin (KMO) calculation used to estimate the size shows the number .847 (KMO> .5), thereby indicating that the sample is adequate. Additionally, Bartlett's Test of Sphericity was valued at 2,174.746 with a significance of .000 (p <.01). Bartlett's Test and KMO results show that factor analysis can be conducted with the distribution of loading factors for each item shown in **Table 5**.

Table 5. The Distribution of Factor Loading

No.	Factor								
	1	2	3	4	5	6	7	8	9
1	.40	-	.20	-	.42	-	-.59	-	-
2	.54	-	-	-	-	-	-.34	-	-
3	.46	-	-	.52	-	-	-	-	-
4	.40	-	-	.51	-	.22	-	.21	-
5	.58	-	-.29	.24	-	-	-	-	-
6	.38	-	.55	-	-	-	.24	.28	.25

	1	2	3	4	5	6	7	8	9
7	.40	-	.62	-	.35	-	-	-	-
8	.34	.27	-	.40	-	.26	-	-.49	-
9	.49	-.33	-.20	-	-	-	-	-	-
10	.49	-.21	-.22	-	.33	-	.36	-	-
11	.55	-.46	-	-	-	-	-	.24	-
12	.48	-	-	-	-	-.42	.32	-	-
13	.51	-	-	-	-	-.36	-	-.34	-
14	.61	-	-	-	-	-.24	-	-	.26
15	.67	-	-.22	-	.22	-.27	-.21	-	-
16	.57	-	-	-.26	-	-.24	-	-.35	-
17	.34	.39	-	-	-	-.23	.20	.21	-.39
18	.51	-.27	-	-	-	-	-	-	.21
19	.55	-.27	-	-	-	-	-.28	-	-
20	.23	.64	-	-	-	-	-	.24	-
21	.38	.33	-	-.29	-	-	-	-	.24
22	.59	-	.32	-	-.21	-	.21	-	-.32
23	.50	-	.49	-	-	-	-	-	-.34
24	.60	-	.28	-.22	-	-	-	-	-
25	.24	.56	-	-	-	-.26	-	-	.22
26	.56	-.28	-	-	-.38	-	-	-	-
27	.60	-	-.27	-	-.22	-	-	-	-
28	.36	-	-	-	-.57	-	-	-	-
29	.55	-.34	-	-	-	-	-	-	.25
30	.38	.30	-	-.40	.29	.41	-	-	-
31	.35	.36	-	-.29	-	.35	-	-	.27
32	.51	.36	-	-	-.27	-	-	.28	-

Table 5 shows that 32 items measuring IMPI-D32 have a dominant loading factor, proving it is a unidimensional scale construct.

Score Categorization

Calculation of hypothetical score categorization is conducted by calculating the hypothetical mean (μ) and standard deviation (σ) first. The hypothetical mean (μ) is calculated as follows:

$$\mu = \frac{Nt + Nr}{2} k$$

$$\mu = \frac{5 + 1}{2} (32)$$

$$\mu = 96$$

Where: k = item total
          N_t = maximum score
          N_r = minimum score

Calculation of hypothetical standard deviation (σ) is as follows: $\sigma = \frac{Nt - Nr}{6} k$

$$\sigma = \frac{5 - 1}{6} 32$$

$$\sigma = \frac{1}{6}(128)$$

$$\sigma = 21$$

The next stage divides scores into five categories based on the hypothetical mean (μ) and hypothetical standard deviation (σ), as shown in **Table 6**.

Table 6. Score Categorization

Category	Score Category	Interpretation
$X \leq (\mu - 1.5\,\sigma)$	$X \leq 64$	Very Low
$(\mu - 1.5\,\sigma) < X \leq (\mu - .5\,\sigma)$	$64 < X \leq 85$	Low
$(\mu - .5\,\sigma) < X \leq (\mu + .5\,\sigma)$	$85 < X \leq 107$	Average
$(\mu + .5\,\sigma) < X \leq (\mu + 1.5\,\sigma)$	$107 < X \leq 128$	High
$(\mu + 1.5\,\sigma) < X$	$128 < X$	Very High

Discipline is one of the important characteristics that need to be possessed by Indonesian Air Force personnel in an effort to achieve national defense goals. Personnel discipline is defined as a character that arises due to individuals' awareness, self-will, and initiative to carry out applicable rules and the capacity to override and resist temptation. Hurlock (1993) stated that discipline is a terminology that comes from the word disciple, which means voluntarily following a leader or authority. The existence of awareness, self-initiative, and volunteerism distinguishes the discipline in this study from work disciplines that are widely discussed in the realm of industrial and organizational psychology. Therefore, two existing psychological constructs were integrated into this research: self-discipline (Duckworth & Seligman, 2006; Duckworth & Kern, 2011; Taylor, Kuo, & Sullivan, 2002; Zimmerman & Kitsantas, 2014) and motivation to obey (Spector, 1978).

The differences between the two concepts formulate the accuracy of the constructs used in this study and their suitability in measuring the instrument on the target population. The concepts of general discipline and military are common in emphasizing behavior according to applicable rules. The discipline meant by the two contexts is internalized self-discipline, which is associated with understanding and implementing rules as a form of respect and self-awareness to achieve common goals.

Discipline in a general context emphasizes the emergence of disciplinary behavior in the absence of external motivation and encouragement. Individuals who possess self-discipline are shown by their absolute motivation to behave in a disciplined manner. Meanwhile, discipline in the military context is an absolute requirement for a soldier, which includes punishment for those who commit indisciplinary actions. The presence of punishment or sanctions in the military is a precautionary measure and not an external threat or motivation for soldiers to conduct discipline.

There are five basics that affect military discipline both internally and externally, namely personality, social background, leader, work rules, and other situations. Meanwhile, the general concept of discipline, which is explained by self-discipline, assesses it as a basic internal character of soldiers.

In scale construction, the readability test involved Air Force TNI personnel as a data source. However, in this test, many items are considered difficult to understand. The results of the discussions carried out stated that most items were problematic because the terms used were inappropriate or not commonly used in the context of the Air Force. The term 'rigid rules that cause frustration' is difficult to understand because it is usually associated with strict rules. The term frustration used is interpreted as a state of despair, while what arises from these strict rules is annoyance and nagging. The term used in measuring instruments must be better adapted to the habits that exist in the military environment, especially in the Indonesian Air Force. Adjusting the use of terms with military culture is important in view of the diverse educational backgrounds of warriors and the ability to understand different languages. In addition, the term adjustment is also carried out to make measuring devices that have good psychometric properties by avoiding the possibility of participants responding by guessing or carelessly.

The number of items specified in the blueprint is 32, which indicates the correction of the spurious overlap effect as a measuring tool with more than 30 items does not need to be corrected (Guilford, 1956; Azwar, 2012). Azwar's (2012) technique was used to produce a more accurate estimate based on the power of discrimination, which uses a corrected item-total correlation (Azwar, 2012).

IMPI-D32 has the ability to distinguish between soldiers with low and high discipline, as indicated by the good item discrimination. Furthermore, the discriminatory power of the measuring instrument is indicated

by a mean value of .407. This illustrates the ability to measure instruments to distinguish individuals who have high and low discipline quite well.

The reliability estimation of the results produced by the IMPI-D32 measuring instrument is conducted using the Spearman-Brown formula, which provides a more stable estimate. The IMPI-D32 measuring device also fulfills the assumption that parallel tests are indicated by each hemisphere's mean and variance scores with equal and indifferent correlation.

The reliability estimation of the IMPI-D32 measurement results shows a reliability value of 0.916. This amount is satisfactory because it has exceeded the initial minimum acceptance limit of the reliability measurement results (Anastasi, 1976; Cohen & Swerdlikl, 2013). Although the IMPI-D32 score is satisfactory, it enables individuals to make important decisions as a basic diagnostic using the measuring instrument and other devices (Azwar, 2017). The reliability coefficient of .916 is estimated to determine the IMPI-D32 measuring instrument with a 91% pure score. Overall, this shows that the soldiers are disciplined with a reliable estimated standard error of measurement.

The standard error of measurement is used to estimate individual measurement changes when repeated tests are conducted. Therefore, the higher the reliability coefficient, the smaller the standard error of measurement using the IMPI-D32 gauge at 3.199, which is within the X score of $\pm$ 6. Large pure score intervals can have an impact on individual scores in the threshold category. When the individual score is 111 and included in the category of high discipline, the range of pure individual A score becomes 105-117. Meanwhile, scores which lower than 107 are categorized as a medium discipline with risks in interpreting categories that occur inaccurately. The standard error of measurement results in an estimated range of pure scores of $X \pm 6$ is relatively high because it contains 128 points. When the range of pure scores is compared to the measuring instruments, the distance $X \pm 6$ is not a distance that is not too wide.

In estimating the construct validity, the IMPI-D32 measuring device uses a unidimensional construct, indicating one dominant factor. Dimensions are defined as the number of equal factors or latent constructs required to form relationships between measurement variables (Netemeyer, Bearden, & Sharma, 2003). **Table 5** shows some items loading on other greater factors. However, these items still have a charge that cannot be ignored on the dominant factor due to the psychological scale.

The categorization conducted in this study is hypothetical because the sample size is not large enough to make empirical norms (n = 229). This is similar to the research conducted by McGrew, LaForte, and Schrank (1994) in making empirical norms for the Woodcock-Johnson IV test using a sample of more than 7,000 people. The risk of using hypothetical categorization of research scores is the proportion of score distribution in accordance with the normal curve.

Conclusion

In conclusion, this study successfully designed an IMPI-D32 tool capable of measuring the discipline level of the TNI Air Force soldiers based on their psychometric properties and standards. This tool can also be used to measure the discipline of TNI Air Force personnel at the level of First Perwira and Bintara.

Future studies can be conducted regarding ways to develop the IMPI-D32 measuring instrument to determine the size of the data source and the hypothetical categorization used in this study for a more stable empirical categorization. Subsequent development can also be related to the predictive validity of the IMPI-D32 due to the repeated use of the discipline data as predictive criteria. The development of predictive validity can be conducted by remembering that the IMPI-D32 reveals discipline as an actual variable.

Studies can also be conducted to estimate the external validity of the IMPI-D32 by analyzing the criteria for other measuring tools to determine the external validity needed for high-impact decision-making. Another suggestion that can be made is related to the modification of measuring instruments by adding items that represent discipline in special situations or corps for soldiers serving in the Special Forces (PASKHAS) or those to be deployed in certain missions. Studies can also be conducted on developing data sources on the Air Force and other TNI such as the Army and Navy, given the similarity of the military, and cultural backgrounds in Indonesia.

Conflict of Interest

There is no conflict of interest in the research. The Indonesian version scale is available by request to the corresponding author email.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, *45*, 131–142.

Anastasi, A. (1976). *Psychological Testing* (4th ed.). New York, NY, US: Macmillan Publishing.

Azwar, S. (2012). *Reliabilitas dan validitas (Ed. 4)*. Yogyakarta: Pustaka Pelajar.

Azwar, S. (2017). *Penyusunan skala psikologi*. Yogyakarta: Pustaka Pelajar.

Cohen, R. J., & Swerdlikl, M. E. (2013). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (7th ed.). New York, NY, US: McGraw-Hill.

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality*, *45*, 259–268. https://doi.org/10.1016/j.jrp.2011.02.004

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge : Gender in self-discipline , grades , and achievement test scores. *Journal of Educational Psychology*. *98*(1), 198–208. https://doi.org/10.1037/0022-0663.98.1.198

Guilford, J. (1956). *Fundamental statistics in psychology and education* (3 ed.). New York, NY: McGraww Hill.

Hartmann, E., Sunde, T., Kristensen, W., & Martinussen, M. (2003). Psychological measures as predictors of military training performance. Journal of Personality Assessment, 80(1), 87–98. https://doi.org/10.1207/S15327752JPA8001_17

Hurlock, Elizabeth B. (1993). Psikologi Perkembangan: Suatu pendekatan sepanjang rentang kehidupan (edisi kelima). Jakarta: Erlangga.

McGrew, K., LaForte, E., & Schrank, F. (1994). *Woodcock-Johnson IV technical manual*. Riverside: The Riverside Publishing Company.

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2017). A closer look at the Test of Personal Intelligence (TOPI). *Personality and Individual Differences*, *111*, 301–311. https://doi.org/10.1016/j.paid.2017.02.008

Mihai, S. C., Corneliu, M., & Sandina, I. (2014). Dimensions of the personality of the military students from Land Forces Academy of Sibiu: A psychometric approach. Procedia-Social and Behavioral Sciences, 127, 479–483. https://doi.org/10.1016/j.sbspro.2014.03.294

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks, Callifornia: Sage Publication, Inc.

Spector, B. I. (1978). *Military self-discipline: A motivational Analysis*. Arlington, Virginia.

Taylor, A. F., Kuo, F. E., & Sullivan, W. C. (2002). Views of nature and self-discipline: Evidence from inner city children. *Journal of Environmental Psychology*, *22*, 49–63. https://doi.org/10.1006/jevp.2001.0241

Thorndike, R. M. (1997). The early history of intelligence testing. In *Contemporary intellectual assessment: Theories, tests, and issues.* (pp. 3–16). New York, NY, US: Guilford Press.

Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students ' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology*. *39*, 145–155.