

Validity of the Dark Triad Dirty Dozen (DTDD) test-Indonesian version

Lalita Devi¹, Lina Natalya¹, Ide Bagus Siaputra^{1*}, Peter Karl Jonason^{2,3}

¹Faculty of Psychology, Universitas Surabaya, Indonesia

²Institute of Psychology, University of Cardinal Stefan Wyszyński, Poland

³Department of General Psychology, University of Padua, Italy

Corresponding author: std@staff.ubaya.ac.id

ARTICLE INFO

Article history

Received January 21, 2024

Revised July 22, 2024

Accepted July 27, 2024

Keywords

Confirmatory factor analysis

Dark triad dirty dozen

Measurement invariance

Principal component analysis

Validation

ABSTRACT

This study addresses the need for a culturally adapted and validated measurement tool for assessing Dark Triad traits in the Indonesian population. The Dark Triad, consisting of narcissism, Machiavellianism, and psychopathy, are associated with various undesirable social and behavioral outcomes worldwide. Our research objective was to validate the Dark Triad Dirty Dozen scale within the Indonesian setting, using a multi-stage methodology, including Principal Component Analysis, Confirmatory Factor Analysis, reliability testing, and invariance testing. A sample of 429 Indonesian undergraduate students (17-23 years old) completed the adapted Dark Triad Dirty Dozen Scale, revealing strong psychometric properties, including validity by confirming a three-factor structure consistent with the theoretical framework, with model fit indices meeting established criteria (RMSEA = 0.054; GFI = 0.932; AGFI = 0.890; CFI = 0.971; NFI = 0.914; TLI = 0.959) and high reliability ($\alpha \geq 0.7$). There is also measurement invariance across men and women in DTDD. In conclusion, the adapted Dark Triad Dirty Dozen Scale is a valid and reliable means to assess Dark Triad traits in Indonesia, contributing to a deeper understanding of personality traits and their implications in this unique cultural context.

This is an open access article under the CC BY-SA license.



Introduction

According to Paulhus and Williams (2002), dark personality refers to a cluster of traits associated with behavior considered contrary to social norms or socially aversive. In their study, they used the term "Dark Triad" to represent three behavioral indicators that can be grouped under dark personality, namely (1) subclinical psychopathy, (2) subclinical narcissism, and (3) machiavellianism. Subclinical psychopathy involves a lack of empathy, immoral behavior, callousness, and an inability to feel guilty about antisocial actions (Jonason & Webster, 2010). Narcissism entails excessive self-focus, a need for recognition, and a tendency to feel superior to others (Jonason & Webster, 2010). Meanwhile, Machiavellianism involves manipulative tactics such as control, deception, and flattery, as well as a tendency to manipulate others for personal gain (Jonason & Webster, 2010). Since first introduced by Paulhus and Williams (2002), the concept of the Dark Triad has become the focus of increasingly intensive research, as evidenced by recent studies (Furnham et al.,

2013; Muris et al., 2017; Miller et al., 2019), both in attempts to understand its negative implications (e.g. Curtis et al., 2022; Esteves et al., 2021; Zhao et al., 2016) as well as other aspects related to it (e.g. Birkás et al., 2018; Lee et al., 2013).

Based on the searches carried out, it was discovered that in a study conducted by Nguyen et al. (2021) with 447 employee samples, 15.9% exhibited high levels of the dark triad (malevolent). Other results were also found in a study conducted in Indonesia, where 2.7% of individuals from a sample of 75 Civil Servants (PNS) were identified with high levels of the dark triad, despite the majority of the sample (73.3%) falling into the low category (Putri et al., 2021). In contrast, different results were found in another study in Indonesia conducted by Banowati and Nugraha (2022), revealing that 85% of 200 social media users aged 18-25 exhibited very high levels of the Dark Triad traits. This raises concerns about potential negative consequences resulting from deviant behavior associated with this personality trait. Deviant behaviors linked to the Dark Triad traits include corruption (Putri et al., 2021; Zhao et al., 2016), deception (Harrison et al., 2016), academic dishonesty (Curtis et al., 2022; Esteves et al., 2021), cyberbullying (Banowati & Nugraha, 2022; Safaria et al., 2020), as well as counterproductive work behavior (O'Boyle et al., 2012). Therefore, it is crucial to employ accurate measurements of these traits to identify the personality characteristics of individuals precisely, aiding in the investigation of potential negative consequences that may arise from these personality traits.

Several studies on the Dark Triad traits also explore common factors influencing the interaction between the three personality traits of the Dark Triad. Jonason and Luévano (2013) suggest that the factor structure of the Dark Triad traits has a complex relationship, illustrating the concept of a bifactor. Their review highlights a common factor that causes overlap among psychopathy, narcissism, and Machiavellianism. Similar findings are echoed in research by McLarnon and Tarraf (2017), stating that the Dark Triad traits is a complex and closely related personality phenomenon, making the identification of common factors crucial in understanding the relationship between these three personality traits. Common characteristics identified in previous research include callousness (Jones & Figueredo, 2013; Paulhus, 2014), manipulateness (Jones & Figueredo, 2013), low agreeableness (Jakobwitz & Egan, 2006), or low levels of honesty-humility (Lee & Ashton, 2005). These findings raise questions about whether the Dark Triad traits can be assessed as a single factor or as three distinct yet interrelated aspects (psychopathy, narcissism, and Machiavellianism) (McLarnon & Tarraf, 2017). Therefore, understanding the existence of common factors becomes relevant in validating measures of the Dark Triad traits.

The concept of the Dark Triad traits has been measured in several studies in Indonesia using the Short Dark Triad (SD3) developed by Jones and Paulhus (2014). Meanwhile, there is a shorter measurement tool from Jonason and Webster (2010) known as the Dark Triad "Dirty Dozen" (DTDD), consisting of a total of 12 items compared to SD3, which comprises a total of 27 items. The use of a shorter measurement tool can offer benefits in terms of time and respondent energy efficiency (Jonason & Webster, 2010). However, it is essential to note that this efficiency must be accompanied by high validity standards (Smith et al., 2000). A brief measurement tool should be able to capture all parts of the construct without omitting essential components of the measured construct (Maples et al., 2014). Therefore, before suggesting the use of a shorter measurement tool to assess the dark triad, its validity needs to be tested first to ensure its accurate application in Indonesian samples.

DTDD also demonstrates superiority in terms of its structure, appearing to be more stable across various languages and cultural contexts compared to SD3 (Rogoza et al., 2021). Furthermore, DTDD itself has been validated in several languages, including Polish (Czarna et al., 2016), Turkish (Özsoy et al., 2017), French-Canadian (Savard et al., 2017), Bengali

(Ahmed et al., 2020), and Korean (Cho et al., 2022). This makes the DTDD measurement tool more practical for research and field use that requires a quick and accurate assessment of the Dark Triad. The validity of DTDD in Indonesia is crucial, considering the collective culture of Indonesian society compared to previous validation research dominated by countries with a more individualistic culture. Because culture can shape the meaning of a construct, the validation of the DTDD scale must be conducted before its use in different cultures (Ma et al., 2021). In this context, this dark personality measurement tool's validation efforts ensure measurement accuracy and aim to understand the concept within the unique socio-cultural dynamics of Indonesia.

Based on previous research, the validation of the DTDD scale has reported differing factor structure results. Some validation studies conducted using confirmatory factor analysis (CFA) found results consistent with the initial structure, where the best model was a three-factor correlated model (e.g., Dinić et al., 2018; El Keshky, 2022; Pineda et al., 2020). On the other hand, different results were found in other studies where the best model was a bifactor model with three factors (e.g., Czarna et al., 2016; Jonason & Luévano, 2013; Savard et al., 2017). Additionally, Cho et al. (2022) found another differing result, with the best model being a bifactor model with two factors, where psychopathy and Machiavellianism were combined into the first factor and narcissism as the second factor. These results leads to differing perspectives regarding scoring. Specifically, there are two conflicting perspectives: the unification approach, where the Dark Triad traits are mapped onto a single common factor, and the separatist approach, which views the three dark triad traits as separate factors (McLarnon & Tarraf, 2021). Therefore, testing the appropriate factor structure for use in Indonesian samples is crucial to ensure scoring the Dark Triad traits as either three interrelated sub-scales or as a single scale.

These different findings provide room for refinement and further understanding of the internal structure of the Dark Triad. Therefore, this validation study aims to determine the appropriate factor structure for the DTDD in the Indonesian sample and ensure consistent measurement. This ensures the tool can be used to identify dark personality characteristics accurately and precisely to aid in researching potential negative consequences associated with this personality trait. Additionally, this study also aims to estimate the reliability of the scale and evaluate measurement invariance based on gender to ensure the DTDD can measure the Dark Triad equally and accurately without gender bias. As for the determination of norms, it is also a crucial aspect in the context of dark personality assessment in Indonesia. Appropriate norms will provide clear guidelines for interpreting assessment results. With norms, assessments become more contextual and relevant, allowing for the optimal use of the DTDD measurement tool in analyzing dark personality in Indonesia.

In conclusion, this research endeavors to deepen our understanding of the Dark Triad within the Indonesian cultural framework, extending beyond psychometric properties and scoring. This study seeks to illuminate the nuances of Dark Triad traits within the Indonesian context, offering insights that could help to better understanding of these traits on individuals and society as a whole.

Method

Research Design

This study is divided into two stages. The first stage involves testing the validity of DTDD using sources of validity evidence from Principal Component Analysis (PCA) and internal consistency. The subsequent stage aims to ensure the validity of DTDD by employing

Confirmatory Factor Analysis (CFA) and assessing internal consistency. Data collection methods involve non-random sampling techniques, specifically accidental sampling. This sampling technique is chosen to ensure diverse sample, thus leading to a more comprehensive assessment of its validity. Participants willing to be involved are requested to complete an informed consent form and a questionnaire through Google Form. In determining the sample size, the researcher follows the rule of thumb guidelines to determine the minimum sample size required for a particular analytical model. The rule of thumb used in determining the sample size suggests ten times the number of items in the measurement tool being validated (Bujang et al., 2022). This study employs the Dark Triad Dirty Dozen (DTDD) measurement tool developed by Jonason and Webster (2010), consisting of a total of 12 items. Therefore, the researcher determines the target sample size for the study to be a minimum of 120 individuals.

Participants

Study 1

Participants in this study consisted of 268 students aged 17-23 years ($M = 20.19$, $SD = 1.15$), comprising 90 men (33.58%), 173 women (64.55%), and five individuals who chose not to disclose their gender (1.87%). The evidence sources used included evidence based on internal structure using Principal Component Analysis (PCA) to examine the grouping of DTDD items, along with reliability analysis to test the internal consistency of DTDD. The time frame for study 1 was around July 2022-January 2023.

Study 2

Participants in this study were 161 students aged 17-23 years ($M = 19.36$, $SD = 1.05$), consisting of 44 men (27.32%), 112 women (69.56%), and five individuals who chose not to disclose their gender (3.10%). The evidence sources used included evidence based on internal structure with Confirmatory Factor Analysis (CFA) to confirm the factor model with the best fit and reliability analysis to test internal consistency. The time frame for study 2 is around June 2023 until August 2023.

Instruments

The measurement tool used in this study is the Dark Triad Dirty Dozen (DTDD), which has been adapted into Bahasa Indonesia (Afifah et al., 2023). DTDD was developed by Jonason and Webster (2010) to measure psychopathy, narcissism, and Machiavellianism. DTDD comprises a total of 12 items, with four items for each dimension. The response scale uses a Likert scale ranging from one (Strongly Disagree) to six (Strongly Agree). Jonason and Webster (2010) reported high reliability for the DTDD scale as follows: $DT = 0.86$, $M = 0.79$, $P = 0.77$, and $N = 0.84$. Similarly, Afifah et al. (2023) found favorable reliability estimates for the DTDD-Indonesian version: $M = 0.768$, $P = 0.759$, and $N = 0.866$.

Table 1

Blueprint Dark Triad Dirty Dozen Scale Indonesian Version

Dimension	Item Number	Total	Examples of Translated Items (in Bahasa Indonesia)
Machiavellianism	1, 2, 3, 4	4	Saya menggunakan pujian untuk memenuhi keinginan saya.
Psychopathy	5, 6, 7, 8	4	Saya cenderung tidak berperasaan atau tidak peka.
Narcissism	9, 10, 11, 12	4	Saya cenderung mencari gengsi atau status.
Total		12	

Meanwhile, behavioral indicators for each aspect of the Indonesian version of DTDD are as follows:

Table 2
Behavioral indicators of DTDD Indonesian Version

Dimension	Definition	Behavioral Indicators
Machiavellianism	Tendency to use other people to fulfill personal desires	Controlling Deceptive Flattering Exploiting others
Psychopathy	Tendency to be not empathetic/behaving anti-socially	Lack of remorse Low morals Callous Apathetic
Narcissism	Tendency to feel entitled	Want to be admired Want to be noticed Want to be appreciated Want special treatment

Data Analysis

Based on the standards outlined in the Standards for Educational and Psychological Testing American Educational Research Association (AERA et al., 2014), five sources of evidence can be used to assess the validity of a psychological test instrument. These sources include evidence based on test content, relationships with other variables, internal structure, response processes, and test consequences. The importance of each validity evidence source varies depending on factors such as the measured construct, the intended use of the test scores, and the assessed population (Reynolds & Livingston, 2014). In this study, the validation testing is conducted using evidence based on internal structure to ensure the most appropriate design for the DTDD instrument, which can accurately and consistently measure dark personality in the Indonesian population.

Factor analysis is one method used to examine a test's internal structure to identify the number of factors or dimensions that constitute a test (Reynolds & Livingston, 2014). There are two types of fact or analysis: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) (Reynolds & Livingston, 2014). EFA is used to explore the representation of items in several latent constructs (Shadiqi, 2023). On the other hand, CFA is used to confirm the alignment of items with the components/dimensions/factors that have been designed (Natalya & Purwanto, 2018). Both types of factor analysis can be used together as a source of internal structure evidence by examining whether the actual structure from the data is consistent with the hypothesized structure (Brown, 2015).

This study uses both EFA with Principal Component Analysis (PCA) and CFA simultaneously to verify whether all items of the Indonesian version of DTDD have been accurately grouped using double-check analysis (Natalya & Purwanto, 2018). CFA is a statistical analysis method within structural equation modeling (SEM) used to verify the number of factors and item grouping while examining the pattern of item-factor relationships (factor loadings) (Brown, 2015). In CFA, several statistical criteria, known as model-fit indices or goodness-of-fit indices, are employed to evaluate the fit between the actual factor structure and the hypothesized factor structure (Reynolds & Livingston, 2014):

Table 3
Cut off value for the goodness of fit indices

Fit Indices	Cut off value	
	Enough	Good
p -value of χ^2	≥ 0.01 (Hu & Bentler, 1999)	> 0.05 (Hu & Bentler, 1999)
TLI	≥ 0.95 (Schermelleh-Engel et al., 2003)	≥ 0.97 (Schermelleh-Engel et al., 2003)
NFI	> 0.90 (Natalya & Purwanto, 2018)	≥ 0.95 (Schermelleh-Engel et al., 2003)
CFI	> 0.90 (Natalya & Purwanto, 2018)	> 0.95 (Hu & Bentler, 1999)
GFI	≥ 0.90 (Natalya & Purwanto, 2018)	≥ 0.95 (Schermelleh-Engel et al., 2003)
AGFI	≥ 0.85 (Natalya & Purwanto, 2018)	≥ 0.90 (Schermelleh-Engel et al., 2003)
RMSEA	< 0.10 (Hu & Bentler, 1999)	< 0.05 (Hu & Bentler, 1999)

In addition to assessing model accuracy, a valid measurement should also have good discriminant validity (Natalya & Purwanto, 2018). Composite Reliability (CR) and Average Variance Extracted (AVE) are two criteria commonly used to support evidence of internal structure validity (Netemeyer et al., 2003). A good CR value is above or equal to 0.7 ($CR \geq 0.7$) (Hair et al., 2014). Meanwhile, for AVE, a good value is above or equal to 0.5 ($AVE \geq 0.5$) (Fornell & Larcker, 1981).

In addition to factor analysis, reliability testing is necessary to support validity testing (Hinton et al., 2014). Reliability testing aims to examine the internal consistency of a measurement instrument. A measurement tool is considered reliable if its measurements are stable and consistent (Reynolds & Livingston, 2014). Generally, a measurement tool is deemed to have satisfactory internal consistency if it has a Cronbach's alpha reliability coefficient of ≥ 0.70 (Hinton et al., 2014).

The purpose of measurement invariance testing is to ensure that there is no potential measurement bias that may arise due to differences between groups, thus ensuring that the analysis results are reliable and have a strong scientific foundation (Chen, 2007). In general, measurement invariance testing involves configural invariance (equality of the number of factors and item composition between groups), metric invariance (equality of factor loadings for each item between groups), and scalar invariance (equality of factor loadings and intercepts for each item between groups) (Muttaqin & Ripa, 2021). Multi-group analysis is conducted following the guidelines proposed by Chen (2007), where differences in CFI coefficients below -0.005 (≤ -0.005) and differences in RMSEA coefficients below 0.010 (≤ 0.010) can be considered as indicators that measurement invariance is acceptable.

Results

Study 1

The validity testing began with conducting PCA to observe the grouping of the 12 items of the Indonesian version of DTDD. The PCA results proposed three factors based on the latent root criterion, a priori criterion, percentage of variance explained criterion, and scree test criterion. All DTDD items had factor loadings above 0.4, and each item was grouped according to its respective dimension, except for item M3, which showed cross-loading with items in the narcissism dimension. However, the factor loading value of item M3 in the Machiavellianism dimension was still higher than its factor loading value in the narcissism dimension (FL Machiavellianism = 0.507 > FL narcissism = 0.473).

Table 4
Factor loading results of the Indonesian version of DTDD items

Dimension	Item	Component			CR	AVE
		Narcissism	Machiavellianism	Psychopathy		
Narcissism	N1	0.808			0.828	0.549
	N2	0.805				
	N4	0.722				
	N3	0.610				
Machiavellianism	M1		0.812		0.798	0.503
	M2		0.755			
	M4		0.725			
	M3	0.473	0.507			
Psychopathy	P3			0.756	0.687	0.363
	P2			0.650		
	P1			0.482		
	P4			0.477		

Study 2

The CFA results for the Indonesian version of DTDD are shown in the Figure 1, Figure 2 and Figure 3. A summary of the results can be seen in Table 5.

Figure 1
Three Factor Correlated Model

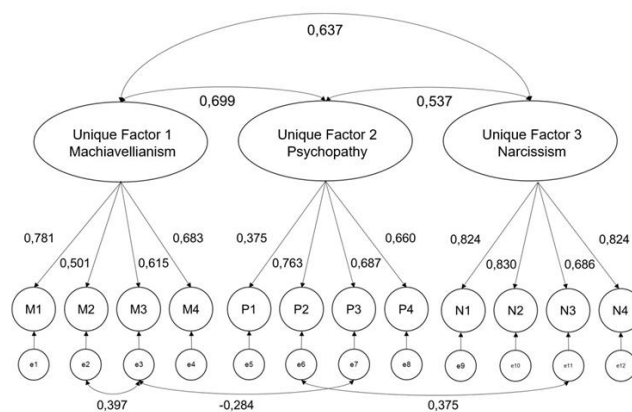


Figure 2
Second Order Model

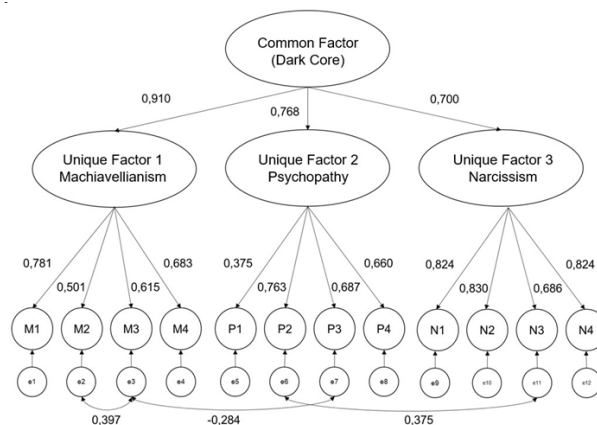


Figure 3
Bifactor Model

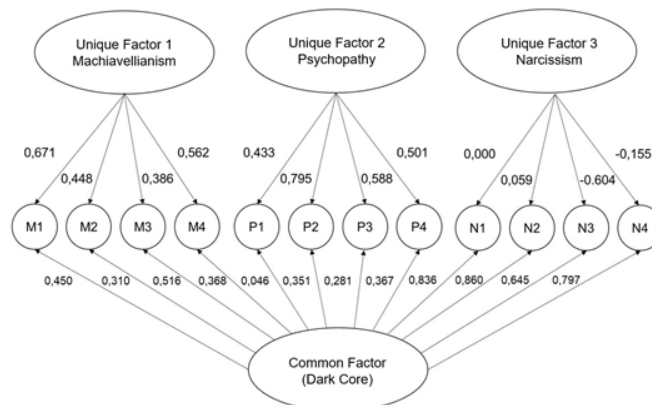


Table 5
Confirmatory Factor Analysis Result of DTDD Indonesian Version

Model	<i>p</i> -value of χ^2	TLI	NFI	CFI	GFI	AGFI	RMSEA
Three-factor correlated model	0.020	0.959	0.914	0.971	0.932	0.890	0.054
Second order model	0.020	0.959	0.914	0.971	0.932	0.890	0.054
Bifactor model	0.008	0.945	0.923	0.968	0.943	0.885	0.062

Note: χ^2 = chi-square, TLI = Tucker–Lewis index, NFI = Normed Fit Index, CFI = Comparative Fit Index, GFI = Goodness of Fit Index, AGFI = Adjusted Goodness of Fit, SRMR = Standardized Root Mean Square Residual, RMSEA = Root Mean Square Error of Approximation.

The CFA testing was conducted on three models: the three-factor correlated model, the hierarchical model, and the bifactor model. Modification indices (MI) were then used to improve the model fit. The testing results for the first model, the three-factor correlated model, showed reasonably adequate results based on the goodness-of-fit criteria. Although overall, all three models had marginal fit indices, the bifactor model did not meet the acceptance criteria for two fit indices, TLI, and *p*-value of χ^2 , which are among the absolute fit measures (Khairi et al., 2021). The bifactor model is not recommended for use based on overall fit indices and specific factor loadings. The presence of a negative correlation (-.604) and a small correlation (.000) suggests the existence of unique factors that remain unexplained. Overall, the three-factor correlated model and the second-order model had better goodness-of-fit values compared to the bifactor model. Therefore, it can be recommended that the hierarchical model is more optimal in explaining the internal structure of the Indonesian version of DTDD. In addition to achieving results that meet the acceptance criteria for the main fit indices, which indicates good model accuracy and internal structural validity, it conceptually provides a clearer explanation of the shared factor across the three traits (Machiavellianism, Psychopathy, and Narcissism).

Table 6
Reliability Analysis Result

No	Dimension	Study 1 (N=268)	Study 2 (N=161)	Number of Items
		α	α	
1	DTDD (Total)	0.736	0.863	12
2	Machiavellianism	0.716	0.762	4
3	Psychopathy	0.430	0.721	4
4	Narcissism	0.760	0.868	4

The reliability testing results for data from study 1 (N = 268) indicate that overall, the DTDD items have good internal consistency with a Cronbach's alpha value above 0.7 ($\alpha = 0.736$). However, when evaluated separately, from the three dimensions of DTDD, the psychopathy dimension in Study 1 does not have good internal consistency with a Cronbach's alpha value below 0.7 ($\alpha = 0.430$). This means that based on the reliability testing in study 1, the items in the psychopathy dimension have less than optimal internal consistency and require further review.

On the other hand, the reliability testing results for data from Study 2 (N = 161) show that overall, and for each dimension, DTDD has Cronbach's alpha that meet the criteria for being considered reliable. Although the first study indicated less satisfactory reliability levels, the internal consistency values in the second study suggest that changes in subject groups can have a significant impact on the reliability of the measurement instrument. Therefore, it can be concluded that the DTDD items have sufficiently reliable internal consistency overall.

Table 7

Gender Measurement Invariance Testing of the Indonesian Version of DTDD

	Model Fit Indices				Model Comparison	
	df	p-value of χ^2	CFI	RMSEA	Δ CFI	Δ RMSEA
Three factor correlation model						
Configural invariance	96	0.001	0.934	0.058		
Metric invariance	105	0.001	0.934	0.056	0.000	-0.002
Scalar invariance	114	0.001	0.930	0.055	-0.004	-0.001
Second order model						
Configural invariance	96	0.001	0.934	0.058		
Metric invariance	105	0.001	0.934	0.056	0.000	-0.002
Scalar invariance	118	0.001	0.929	0.054	-0.005	-0.002
Bifactor model						
Configural invariance	78	0.042	0.970	0.044		
Metric invariance	98	0.021	0.960	0.045	-0.010	0.001
Scalar invariance	107	0.032	0.962	0.042	0.002	-0.003

From the results of the multi-group analysis presented in Table 7, it appears that there are differences in the CFI coefficients below -0.005 (≤ -0.005) and differences in the RMSEA coefficients below 0.010 (≤ 0.010) in the testing of metric and scalar invariance of the DTDD measurement instrument. These results indicate that there is measurement invariance in DTDD based on the respondents' gender.

In other words, the factor structure, factor loadings, and intercepts for each DTDD item do not show significant differences between male and female respondent groups. This finding confirms that DTDD can be reliable and consistent in measuring the Dark Triad, both in male and female respondents, making the results considered valid for the Indonesian population.

To complement the Indonesian version of the DTDD scale, in this study, the researchers established norms. With this information, scale users can interpret the implications of their scores with the reference group established. However, it is important to note that these norms are specific to the population used as the sample, namely Indonesian university students.

Table 8
DTDD Norms Indonesian Version

Category	Ideal Norm (total)	Ideal Norm (dimension)	Group Norm			
			Total (T)	Machiavellianism (M)	Psychopathy (P)	Narcissism (N)
Very Low	< 22	< 7.33	≤ 17	≤ 4	≤ 3	≤ 4
Low	22 - 31	7.33 – 10.66	18 – 26	5 - 8	4 – 6	5 – 9
Quite Low	32 - 41	10.67 – 13.99	27 – 35	9 - 12	7 – 9	10 – 13
Quite High	42 - 51	14.00 – 17.32	36 – 45	13 - 16	10 – 13	14 – 18
High	52 - 62	17.33 – 20.66	46 – 54	17 – 20	14 – 16	19 – 22
Very High	≥ 63	≥ 20.67	≥ 55	≥ 21	≥ 17	≥ 23

Note: Mean T = 35.904; SD T = 9.128; Mean M = 12.515; SD M = 4.021; Mean P = 9.814; SD P = 3.442; Mean N = 13,576; SD N = 4.525

Discussion

This study examined the validity of the Indonesian version of the DTDD measurement instrument through evidence of internal structure using PCA and CFA, supported by reliability testing using Cronbach's alpha. The research began by examining the item grouping based on PCA analysis, and the results were then confirmed through CFA by testing three models (correlation, hierarchy, and bifactor).

Based on the results of PCA, the 12 items of the Indonesian version of DTDD showed good grouping according to the initial design dimensions, which are three dimensions with four items each. One item from the Machiavellianism dimension (M3) cross-loaded into the Narcissism dimension. Cross-loading is likely to occur as the dark triad is a collection of three overlapping dimensions (Paulhus & Williams, 2002). Nevertheless, the factor loading value for item M3 is still higher in the Machiavellianism dimension than in the Narcissism dimension, so the researcher decided to retain the item. Thus, it can be concluded that the 12 items of the Indonesian version of DTDD can be grouped into three dimensions as per the initial design.

Next, the CFA results were used to confirm the proposed factor structure from PCA, which is three factors with four items each according to the a priori design. CFA testing results showed that the three-factor correlated model and hierarchical model had similar goodness-of-fit values, and both met the cut-off values for each index. While the bifactor model also showed good model fit, it was not as good as the correlation and hierarchy models. This result is similar to the findings of Pineda et al. (2020); Dinić et al. (2018); and El Keshky (2022), who found the correlation model to be the best for DTDD. In addition, the multi-group analysis revealed that the Indonesian version of DTDD shows measurement invariance by testing metric and scalar invariance. This finding indicates that there is no difference between the number of factors, item composition, and factor loadings between male and female respondent samples (Chen, 2007).

From the CFA results, a negative correlation of -0.284 was found between the error variance of item M3 and P3 in the three-factor correlation model and hierarchical model. This correlation result may reflect differences in the social strategies applied by individuals who show higher levels in one dimension compared to another (Jones & Mueller, 2022). The Machiavellianism dimension is often associated with more strategic, manipulative, and goal-oriented social strategies (Jones & Paulhus, 2011). In contrast, the Psychopathy dimension tends to describe more impulsive, lack of self-control, and antisocial behavior (Jones & Paulhus, 2011). Therefore, individuals with higher scores in the Machiavellianism dimension may tend to use more focused social strategies, while those showing higher levels in the Psychopathy dimension may engage in more impulsive actions in social interactions

(Furnham et al., 2013). These differences in social strategies can explain the negative correlation found between items in the Machiavellianism dimension and items in the Psychopathy dimension.

Based on previous research, it has been found that the three dimensions of the Dark Triad often overlap, indicating a complex relationship between these dimensions (Paulhus & Williams, 2002; Jones & Paulhus, 2014). Paulhus & Williams (2002) as the first who introduce the concept also indicates in their definition that dark triad is three distinct but overlapping construct of socially aversive traits. Correlating errors in SEM requires a solid theoretical justification, suggesting that certain aspects of the constructs are related and might share unique variance (Perry et al., 2015; Landis et al., 2009). Harrington (2009) also stated that measurement errors can be the result of similar phrases in the item. This seem to be the case for item M2 “*I have used deceit or lied to get my way*” and M3 “*I have use flattery to get my way.*” Previous research indicating overlap among the dimensions of the Dark Triad serves as the theoretical justification for correlating errors between items. This approach ensures that the error correlation aligns with both theoretical underpinnings and empirical evidence obtained from the data, maintaining the integrity of the model.

The reliability testing results from Study 1 and Study 2 showed different outcomes where the internal consistency values for the psychopathy dimension were less reliable in the first study but reliable in the second study. The psychopathy dimension had less internal consistency due to Cronbach's alpha values not meeting the reliability. However, low internal consistency can be caused by the limited number of items on the scale (Schweizer, 2011), with some previous studies also finding lower internal consistency values in the psychopathy dimension (e.g., Dinić et al., 2018; El Keshky, 2022; Jonason & Webster, 2010). Each dimension of DTDD consists of only four items, which can result in relatively low internal consistency when analyzed per dimension compared to the overall internal consistency value (Jonason & Webster, 2010). Cronbach's alpha is a coefficient of internal consistency closely related to the number of items analyzed. The greater the number of items analyzed within a measurement tool or dimension, the likelihood is that the Cronbach's alpha coefficient will increase (Reynolds & Livingston, 2014).

Furthermore, the creation of a shorter scale aims to allow the instrument to measure all aspects of a construct with non-redundant items (Jonason & Webster, 2010). This results in a measurement tool with a limited number of items and tends to be diverse, so the correlation values between items on a shorter scale tend to be low (Rammstedt & Beierlein, 2014). Therefore, lower internal consistency than a longer scale is not a problem if the validity of the scale in terms of content coverage can be maintained (Rammstedt & Beierlein, 2014). Thus, it can be concluded that the Indonesian version of DTDD can consistently measure dark triad tendencies.

Moreover, based on the findings from CR and AVE analysis supporting evidence of internal structure validity, the CR values were above 0.7 for the Machiavellianism and narcissism dimensions ($CR \geq 0.70$). Similarly, only the Machiavellianism and narcissism dimensions had AVE values above 0.5. According to Fornell & Larcker (1981), an AVE value below 0.5 indicates a higher level of average error than the variance captured by the construct. However, if the AVE value is close to 0.5 and other validity criteria are met, the AVE value alone is not enough to indicate a problem. All aspects of DTDD also had CR values in the range of 0.60 - 0.70, which is acceptable according to (Hair et al., 2014). This indicates that all aspects of DTDD are consistent enough. Because overall CFA testing results showed a good model fit for the correlation model according to the initial design and PCA results, it can be said that the Indonesian version of DTDD is valid.

The implications of this research at the theoretical level are that the Indonesian version of DTDD can be considered a valid and reliable measurement tool for assessing the dark triad. Practically, these findings provide a solid foundation for the implementation and interpretation of DTDD in the Indonesian context. Therefore, this research not only contributes to understanding the dark triad concept but also offers useful guidance for practitioners and researchers in accurately applying and interpreting the DTDD scale in the Indonesian population.

However, it is crucial to acknowledge a limitation concerning the need for additional sources of validity evidence. While this study provides strong evidence supporting the internal structure validity of the Indonesian DTDD, future research could benefit from incorporating other sources of validity evidence such as evidence based on test content and correlation with other variables to further validate the scale's effectiveness in capturing dark triad traits.

Conclusion

The results obtained indicate that the Indonesian version of DTDD can be used to measure the dark triad validly and reliably. The first study showed that DTDD could accurately measure the three dimensions, namely Machiavellianism, psychopathy, and narcissism, with item grouping according to the initial design. This result was also confirmed in the second study using CFA, which showed the three-factor model as the most suitable models for the Indonesian version of DTDD. This finding is consistent with previous research that found the three-factor correlated model to be the best for DTDD. Based on the findings of this study, it is known that the Indonesian version of DTDD has good and adequate internal structure validity evidence. Furthermore, the results of gender measurement invariance testing also confirmed that the DTDD measurement tool is not gender biased in the context of the Indonesian population. This indicates that the Indonesian version of DTDD can be used fairly and accurately to measure the levels of the dark triad in individuals without distinguishing their gender. In addition, the factor structure testing also provides a strong foundation for deciding the appropriate scoring method for DTDD. Based on the results of PCA and CFA, it can be ensured that DTDD scoring can be done following the proposed three-factor hierarchy. This process helps validate that the three aspects of the dark triad, including Machiavellianism, psychopathy, and narcissism, are not only interrelated but can also be considered as a single dimension summarizing the complexity of the dark triad. Thus, scoring DTDD can be done by looking at scores per dimension for a more specific understanding of the aspects of the dark triad, while the total score provides a holistic overview of the overall tendency of the dark triad in individuals. Therefore, this research provides a comprehensive understanding of the appropriate scoring method for the DTDD scale.

Acknowledgment

We would like to express my sincere appreciation to Universitas Surabaya for granting permission to conduct the research. Heartfelt thanks to the participants who generously contributed their time to complete the survey. Thank you all for your crucial roles in the success of this research.

Declarations

Author contribution. LD, LN, and IBS collectively contributed to the conception and execution of this research study. LD played a role in proposing the initial research idea,

conducting extensive literature reviews, preparing the measurement tools, overseeing data collection, conducting data analysis, and drafting the initial version of the research article. LN and IBS, serving as research supervisors, provided valuable guidance throughout the research process, offering insights into refining the research topic, scrutinizing measurement tools, conducting verification checks, reviewing language, and contributing to the writing process. PKJ provided valuable input on the methodology and validation aspects of the study and also contributed to the review and editing of the manuscript.

Funding statement. The authors themselves funded this research.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

References

- Afifah, A. S. A. N., Lumintu, A., Pangeran, G. A., Sibuea, M. F. M., Zahra, S. R., Shindied, T. K., & Natalya, L. (2023). Dark Triad Dirty Dozen Scale (DTDD) Versi Bahasa Indonesia. [Dark Triad Dirty Dozen Scale (DTDD) Indonesiaian Version]. In L. Natalya & G. Wibaningrum (Eds.), *Measuring Human Behavior: 32 Psychological Tools – A Sourcebook* (1st ed.). Fakultas Psikologi - Universitas Surabaya.
- Ahmed, O., Naher, L., Islam, R., Akter, M., & Deb, S. (2020). Psychometric analyses of the Bangla version of the Dark Triad Dirty Dozen. *Heliyon*, 6(11), e05341. <https://doi.org/10.1016/j.heliyon.2020.e05341>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA.
- Banowati, A., & Nugraha, S. (2022). Pengaruh kepribadian dark triad terhadap perilaku cyberbullying pada pengguna media sosial [The influence of dark triad personality on cyberbullying behavior among social media users]. *Bandung Conference Series: Psychology Science*, 2(3), 682–689. <https://doi.org/10.29313/bcps.v2i3.2879>
- Birkás, B., Pátkai, G., & Csathó, Á. (2020). The mediating role of the dark triad between life history strategy and perceived stress factors. *Psychological Reports*, 123(2), 252–265. <https://doi.org/10.1177/0033294118818095>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Bujang, M. A., Yee, L. K., & Khee, H. Y. (2022). *A step-by-step guide to questionnaire validation research*. Zenodo. <https://doi.org/10.5281/zenodo.6801209>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cho, H., Fengge, L., Ma, H., & Ock, J. (2022). Validation and factor structure analysis of the K-Dirty Dozen. *Korean Journal of Industrial and Organizational Psychology*, 35(2), 299–326. <https://doi.org/10.24230/kjiop.v35i2.299-326>
- Curtis, G. J., Correia, H. M., & Davis, M. C. (2022). Entitlement mediates the relationship between dark triad traits and academic misconduct. *Personality and Individual Differences*, 191, 111563. <https://doi.org/10.1016/j.paid.2022.111563>
- Czarna, A. Z., Jonason, P. K., Dufner, M., & Kossowska, M. (2016). The Dirty Dozen Scale: Validation of a Polish version and extension of the nomological net. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00445>
- Dinić, B. M., Petrović, B., & Jonason, P. K. (2018). Serbian adaptations of the Dark Triad Dirty Dozen (DTDD) and Short Dark Triad (SD3). *Personality and Individual Differences*, 134, 321–328. <https://doi.org/10.1016/j.paid.2018.06.018>

- El Keshky, M. E. S. (2022). Psychometric properties of an Arabic version of the Dark Triad Dirty Dozen Scale. *Australian Journal of Psychology*, 74(1). <https://doi.org/10.1080/00049530.2022.2138543>
- Esteves, G. G. L., Oliveira, L. S., de Andrade, J. M., & Menezes, M. P. (2021). Dark triad predicts academic cheating. *Personality and Individual Differences*, 171, 110513. <https://doi.org/10.1016/j.paid.2020.110513>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The dark triad of personality: A 10-year review. *Social and Personality Psychology Compass*, 7(3), 199–216. <https://doi.org/10.1111/spc3.12018>
- Hair, J. F. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. SAGE.
- Harrington, D. (2009). *Confirmatory factor analysis*. Oxford University Press.
- Harrison, A., Summers, J., & Mennecke, B. (2018). The effects of the dark triad on unethical behavior. *Journal of Business Ethics*, 153(1), 53–77. <https://doi.org/10.1007/s10551-016-3368-3>
- Hinton, P., McMurray, I., & Brownlow, C. (2014). *SPSS Explained*. Routledge. <https://doi.org/10.4324/9781315797298>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jakobowitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual Differences*, 40(2), 331–339. <https://doi.org/10.1016/j.paid.2005.07.006>
- Jonason, P. K., & Luévano, V. X. (2013). Walking the thin line between efficiency and accuracy: Validity and structural properties of the Dirty Dozen. *Personality and Individual Differences*, 55(1), 76–81. <https://doi.org/10.1016/j.paid.2013.02.010>
- Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the Dark Triad. *Psychological Assessment*, 22(2), 420–432. <https://doi.org/10.1037/a0019265>
- Jones, D. N., & Figueredo, A. J. (2013). The core of darkness: Uncovering the heart of the Dark Triad. *European Journal of Personality*, 27(6), 521–531. <https://doi.org/10.1002/per.1893>
- Jones, D. N., & Mueller, S. M. (2022). Is Machiavellianism dead or dormant? The Perils of Researching a Secretive Construct. *Journal of Business Ethics*, 176(3), 535–549. <https://doi.org/10.1007/s10551-020-04708-w>
- Jones, D. N., & Paulhus, D. L. (2011). The role of impulsivity in the Dark Triad of personality. *Personality and Individual Differences*, 51(5), 679–682. <https://doi.org/10.1016/j.paid.2011.04.011>
- Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3). *Assessment*, 21(1), 28–41. <https://doi.org/10.1177/1073191113514105>
- Khairi, M. I., Susanti, D., & Sukono, S. (2021). Study on structural equation modeling for analyzing data. *International Journal of Ethno-Sciences and Education Research*, 1(3), 52–60. <https://doi.org/10.46336/ijeer.v1i3.295>
- Landis, R. S., Edwards, B. D., & Cortina, J. M. (2009). On the practice of allowing correlated residuals among indicators in structural equation models. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends:*

- Doctrine, verity and fable in the organizational and social sciences* (pp. 193–214). Routledge.
- Lee, K., & Ashton, M. C. (2005). Psychopathy, Machiavellianism, and Narcissism in the Five-Factor Model and the HEXACO model of personality structure. *Personality and Individual Differences, 38*(7), 1571–1582. <https://doi.org/10.1016/j.paid.2004.09.016>
- Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Gallucci, A. (2013). Sex, power, and money: Prediction from the dark triad and honesty–humility. *European Journal of Personality, 27*(2), 169–184. <https://doi.org/10.1002/per.1860>
- Ma, G. X., Born, M. P., Petrou, P., & Bakker, A. B. (2021). Bright sides of dark personality? A cross-cultural study on the dark triad and work outcomes. *International Journal of Selection and Assessment, 29*(3–4), 510–518. <https://doi.org/10.1111/ijasa.12342>
- Maples, J. L., Lamkin, J., & Miller, J. D. (2014). A test of two brief measures of the dark triad: The dirty dozen and short dark triad. *Psychological Assessment, 26*(1), 326–331. <https://doi.org/10.1037/a0035084>
- McLarnon, M. J. W., & Tarraf, R. C. (2017). The Dark Triad: Specific or general sources of variance? A bifactor exploratory structural equation modeling approach. *Personality and Individual Differences, 112*, 67–73. <https://doi.org/10.1016/j.paid.2017.02.049>
- McLarnon, M. J. W., & Tarraf, R. C. (2021). Getting to the core: How “(dis)honest” is the core of the Dark Triad? *Personality and Individual Differences, 171*, 110545. <https://doi.org/10.1016/j.paid.2020.110545>
- Miller, J. D., Vize, C., Crowe, M. L., & Lynam, D. R. (2019). A critical appraisal of the dark-triad literature and suggestions for moving forward. *Current Directions in Psychological Science, 28*(4), 353–360. <https://doi.org/10.1177/0963721419838233>
- Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature. *Perspectives on Psychological Science, 12*(2), 183–204. <https://doi.org/10.1177/1745691616666070>
- Muttaqin, D., & Ripa, S. (2021). Psychometric properties of the Indonesian version of the Depression Anxiety Stress Scale: Factor structure, reliability, gender, and age measurement invariance. *Psikohumaniora: Jurnal Penelitian Psikologi, 6*(1), 61–76. <https://doi.org/10.21580/pjpp.v6i1.7815>
- Natalya, L., & Purwanto, C. V. (2018). Exploratory and confirmatory factor analysis of the Academic Motivation Scale (AMS)–Bahasa Indonesia. *Makara Human Behavior Studies in Asia, 22*(1), 29. <https://doi.org/10.7454/hubs.asia.2130118>
- Netemeyer, R., Bearden, W., & Sharma, S. (2003). *Scaling procedures*. SAGE Publications, Inc. <http://dx.doi.org/10.4135/9781412985772>
- Nguyen, N., Pascart, S., & Borteyrou, X. (2021). The dark triad personality traits and work behaviors: A person-centered approach. *Personality and Individual Differences, 170*, 110432. <https://doi.org/10.1016/j.paid.2020.110432>
- O’Boyle, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the Dark Triad and work behavior: A social exchange perspective. *Journal of Applied Psychology, 97*(3), 557–579. <https://doi.org/10.1037/a0025679>
- Özsoy, E., Rauthmann, J. F., Jonason, P. K., & Ardiç, K. (2017). Reliability and validity of the Turkish versions of Dark Triad Dirty Dozen (DTDD-T), Short Dark Triad (SD3-T), and Single Item Narcissism Scale (SINS-T). *Personality and Individual Differences, 117*, 11–14. <https://doi.org/10.1016/j.paid.2017.05.019>
- Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science, 23*(6), 421–426. <https://doi.org/10.1177/0963721414547737>

- Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*(6), 556–563. [https://doi.org/10.1016/S0092-6566\(02\)00505-6](https://doi.org/10.1016/S0092-6566(02)00505-6)
- Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science, 19*(1), 12–21. <https://doi.org/10.1080/1091367X.2014.952370>
- Pineda, D., Sandín, B., & Muris, P. (2020). Psychometrics properties of the Spanish version of two Dark Triad scales: The Dirty Dozen and the Short Dark Triad. *Current Psychology, 39*(5), 1873–1881. <https://doi.org/10.1007/s12144-018-9888-5>
- Putri, K.W., Rahayu, P.Y., & Arunima, A. (2021). Dark triad personality as predictor of corrupt intention on the state civil apparatus. *Journal of Educational, Health and Community Psychology, 10*(2), 375. <https://doi.org/10.12928/jehcp.v10i2.20649>
- Rammstedt, B., & Beierlein, C. (2014). Can't we make it any shorter? *Journal of Individual Differences, 35*(4), 212–220. <https://doi.org/10.1027/1614-0001/a000141>
- Reynolds, C. R., & Livingston, R. B. (2014). *Mastering modern psychological testing: Theory & methods*. Pearson.
- Rogoza, R., Żemojtel-Piotrowska, M., Jonason, P. K., Piotrowski, J., Campbell, K. W., Gebauer, J. E., Maltby, J., Sedikides, C., Adamovic, M., Adams, B. G., Ang, R. P., Ardi, R., Atitsogbe, K. A., Baltatescu, S., Bilić, S., Bodroža, B., Gruneau Brulin, J., Bundhoo Poonosamy, H. Y., Chaleeraktragoon, T., ... Włodarczyk, A. (2021). Structure of Dark Triad Dirty Dozen across eight world regions. *Assessment, 28*(4), 1125–1135. <https://doi.org/10.1177/1073191120922611>
- Safaria, T., Lubabin, F., Purwandari, E., Ratnaningsih, I. Z., Nofrans, M. K., Saputra, E., Ipak, E., Esita, R., Nazriani, D., & Mariyati, L. E. (2020). The role of dark triad personality on cyberbullying: Is it still a problem? *International Journal of Scientific & Technology Research, 9*(2), 4256–4260.
- Savard, C., Simard, C., & Jonason, P. K. (2017). Psychometric properties of the French-Canadian version of the Dark Triad Dirty Dozen. *Personality and Individual Differences, 119*, 122–128. <https://doi.org/10.1016/j.paid.2017.06.044>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74. <https://doi.org/10.23668/psycharchives.12784>
- Schweizer, K. (2011). Some thoughts concerning the recent shift from measures with many items to measures with few items. *European Journal of Psychological Assessment, 27*(2), 71–72. <https://doi.org/10.1027/1015-5759/a000056>
- Shadiqi, M. A. (2023). *Statistik untuk Penelitian Psikologi dengan SPSS* [Statistics for psychological research using SPSS]. Rajawali Pers.
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*(1), 102–111. <https://doi.org/10.1037/1040-3590.12.1.102>
- Zhao, H., Zhang, H., & Xu, Y. (2016). Does the dark triad of personality predict corrupt intention? The mediating role of belief in good luck. *Frontiers in Psychology, 7*(2), 71–72. <https://doi.org/10.3389/fpsyg.2016.00608>