ADAPTING THE DIRTY DOZEN IN THE INDONESIAN CONTEXT: PSYCHOMETRIC EVALUATION OF THE DARK TRIAD TRAITS FOR EDUCATIONAL AND PSYCHOLOGICAL RESEARCH

Abdurosid Nur Ali[1*], Wuri Prasetyawati[2]
[1,2] Universitas Indonesia, Indonesia
*Corresponding Author: abdurosid.nur@ui.ac.id

ABSTRACT
Dark triad personality traits—narcissism, psychopathy, and Machiavellianism—are recognized as important constructs in psychological research but have lacked a concise, psychometrically sound assessment tool in Indonesia. This study aimed to adapt and validate the Dirty Dozen (DD) instrument into Bahasa Indonesia to address this gap. Using a sample of 157 adults recruited online, the research employed a rigorous adaptation process, including translation, back-translation, expert judgment, and pilot testing. Psychometric properties were examined through confirmatory factor analysis (CFA) and reliability analysis. Results indicated that the Indonesian DD demonstrated strong reliability and validity for the narcissism and Machiavellianism subscales, while the psychopathy subscale showed lower internal consistency, likely due to cultural and linguistic challenges in item adaptation. The CFA supported the original three-factor structure, with significant overlap between Machiavellianism and psychopathy, and no gender differences found in any domain. In conclusion, the Indonesian DD is a practical and efficient tool for measuring dark triad traits in local populations, though refinement of psychopathy items and response scale categories is recommended for future research.

Keywords: adaptation, dark triad, Indonesia, measurement, psychometrics

INTRODUCTION

The dark personality, also known as the “dark triad”, has been studied intensively and extensively by several researchers, especially in the field of psychology. The existence of dark personality was first proposed through Jung's personality theory, where Jung stated that every individual has a shadow that individuals often do not realize and do not see it, or even try to hide it from themselves and others, in its development then called the dark triad personality (Paulhus & Williams, 2002). The dark triad personality reveals that every individual has a dark side, so called narcissism, psychopathy, and Machiavellianism. Individuals with narcissism trait have an exaggerated view of their self-worth and a high sense of superiority, focus only on themselves, arrogant, and tend to be exploitative in interpersonal relationships, and view others as a means to fulfill their need for admiration and reinforcement of self-view (Campbell et al., 2000, 2002; Rhodewalt & Peterson, 2009). Individuals with psychopathy trait use destructive patterns, show dysfunctional behavior in their interpersonal relationships, tend not to care about the guilt of their own behavior, high impulsivity, immoral, unable to empathize, often show themselves without guilt, cold, and also insensitive to the feelings of others (Hare, 1999; Williams & Paulhus, 2004).

Individuals with Machiavellianism trait are characterized by deceit, using interpersonal strategies in manipulating others for their personal gain, tend to exploit the weaknesses of others for their personal goals, have a view that it is better to manipulate than to be manipulated by others, and lack of emotional attachment in social interactions (Christie & Geis, 1970; Fehr et al., 1992; Harrel, 1980). The results show that individuals with high scores in these dark triad personality traits will tend to engage in immoral behaviors, such as self-promotion, emotional coldness, duplicity, and aggressiveness. Rauthmann & Kolar (2012) revealed how individuals can perceive the traits of the dark triad in themselves and others by evaluating

based on three criteria, namely desire, consequences for self, and consequences for others. Generally, perceptions of psychopathy and Machiavellianism are close to antagonistic styles characterized by arrogance, calculation and tactics, callousness, and manipulative nature. Narcissism, on the other hand, is perceived to be more unique, due to its desire to be the center of attention, high conscientiousness, and motivation for achievement. So, narcissism is considered better than the other two dimensions. Furnham et al. (2014) revealed that the measurement of these three dark triad traits has its particular concern, especially in the measurement and statistical approaches used in describing the dark triad traits that still seem to overlap but can be distinguished. Then, this study also explained in detail the problems of measurement literature, especially the assessment of the characteristics of manipulative and insensitive individuals. These two attributes are found in each of the dark triad traits, as described in previous research by Rauthmann & Kolar (2012), who stated that psychopathy and Machiavellianism are fairly similar and are concepts that will overlap, as they are difficult to distinguish.

One of the dark triad measurement tools that is often used in studies outside Indonesia is the Dirty Dozen. Studies with the themes of emotion, intelligence, love, perception, politics, religiosity, and behavior at work are often found using the Dirty Dozen measuring instrument (Alper et al., 2021; Jonason & Sherman, 2020; Miao et al., 2019; Szabó et al., 2023; Tajmirriyahi et al., 2021; Zhuang et al., 2022). Meanwhile, dark triad personality research in Indonesia popularly uses the Short Dark Triad (SD3) with a total of 27 items compiled by Jones & Paulhus (2014). Maples et al. (2014) tested two dark triad measurement tools, namely Dirty Dozen (DD) and Short Dark Triad (SD3). The results show that the SD3 is indeed quite consistent in measuring the concept of the dark triad, but in content there are differences with the Dirty Dozen. Narcissism in SD3 only reveals an exaggerated view of the self, whereas narcissism in DD also reveals the vulnerability of the self to threats about its self-worth. Two things that need to be revealed in the construct of narcissism.

Reliability measured on the short dark triad (SD3) adapted for the German population had high and stable test-retest scores (Mach. *rtt*=.81, psycho. *rtt*=.83, narr. *rtt*=.74) on all three factors in the SD3. In the German version of the SD3, the CFA conducted with the three-factor model had valid and fit data. This is also supported when comparing the three-factor model and the one-factor model (Malesza et al., 2019), where the three-factor model is more fit. In the adaptation of SD3 conducted on participants of Serbian nationality, the use of the Item Response Theory (IRT) method to find inter-item correlation produced fit results. The average inter-item correlation on each factor has above-average results (CMI > .30). Although when looking at the results of the analysis per item, there are several items that have factor loading scores and IRT parameters below the average (Dinić et al., 2018). Validity and reliability testing on SD3 that has been adapted to French citizens using B-ESEM ($MD\Delta X^2 = 119.23$; df = 23; $p < .001$) and B-CFA ($MD\Delta X^2 = 217.70$; df = 46; $p < .001$) has fit results for the three-factor model (Gamache et al., 2018). However, in different samples, such as Indonesia, and including dark triad personality as a variable is not followed by clear psychometric properties of the SD3 measuring instrument (Asih & Lutfiyah, 2023; Kaumbur et al., 2018; Nuzulia & Why, 2020; Safaria et al., 2020). Hasanati & Istiqomah (2019) have adapted the Short Dark Triad into Indonesian. The study attempted to show construct validity but did not meet the criteria (misfit). Only 16 items out of 27 items were able to be used but did not meet the rules of adaptation and could not maintain the same number of items and evidence of proper construct validity. The composite internal consistency of the instrument was 0.85, and each domain of Machiavellianism was 0.778, narcissism was 0.73, and psychopathy was 0.823.

Adaptation of the dirty dozen has been widely done in Europe, such as Poland (Czarna et al., 2016), Serbia (Dinić et al., 2018), Turkey (Kayiş et al., 2018), and France (Gamache et al., 2018). One adaptation conducted in Japan by Tamura et al. (2015) showed quite good results with a sample of university students in Japan. The reliability of the total score was 0.81, Machiavellianism was 0.78, psychopathy was 0.55, and narcissism was 0.79, although there were items measuring psychopathy in the psychopathy domain that did not show satisfactory results, possibly due to unequal adaptation results. Reliability alone is not able to

explain the ability of a measuring instrument properly. Evidence of validity also needs to be explained to strengthen the validity of a measurement tool. Therefore, research is needed to adapt other dark triad personality measurement tools that may have better measurement capabilities and clear and qualified psychometric properties. An instrument with a small number of items in quantitative research will be much better than an instrument with a larger number of items. Time and cost effectiveness are key points in using short versions of instruments (Kemper et al., 2019). Kemper et al. (2019) found no difference in the results of using the short and long versions of the instrument in a clinical context, so the advantages and effectiveness of the short version of the measuring instrument should be considered. Other research results that support this argument are research from Akhtar & Sumintono, (2023) who tested the short and long versions of the big five personality trait instrument with the Rasch model approach and still showed good psychometric properties, especially in many response options, item strength, reliability, and item bias. Therefore, the use of the short version is quite comprehensive in measuring a variable. However, the ability of the long version of the instrument can still be used for more precise measurement and better diagnostic capabilities. Researchers attempted to adapt the dirty dozen measurement tool compiled by Jonason & Webster (2010) which amounted to 12 items representing the three dimensions of the dark triad and tested the psychometric properties with classical and modern model testing.

METHOD

Participants

The sample consisted of 157 participants over the age of 17 who were obtained through social media postings and online broadcast messages (M=26.59; SD=7.66) (Table 1). Exclusion criteria were individuals who had not reached the age of 17 and had difficulty reading and understanding sentences in Indonesian. The personal data collected was anonymous and analyzed publicly through an online platform (Google Form). All participants were given and agreed to the online consent form before completing the questionnaire. The minimum number of participants for confirmatory factor analysis is 150 if the data is normally distributed and 265 if it is not normally distributed (Muthén & Muthén, 2002).

Table 1. Characteristic of participants (N=157)

	Demographic	N	Percentage	Cumulative
Gender				
	Female	96	61%	61%
	Male	61	39%	100%
Region (n=154)				
	Java, Bali, & Nusa Tenggara	131	85.06%	85.06%
	Kalimantan	6	3.90%	88.96%
	Sumatra	15	9.74%	98.70%
	Sulawesi, Maluku, & Papua	2	1.30%	100%
Ethnicity				
	Acehnese	1	0.64%	0.64%
	Balinese	1	0.64%	1.27%
	Banjarese	2	1.27%	2.55%
	Bataknese	8	5.10%	7.64%
	Batavians	4	2.55%	10.19%
	Buginese	2	1.27%	11.47%
	Dayak	1	0.64%	12.10%
	Sundanese	33	21.02%	33.12%
	Javanese	80	50.96%	84.08%
	Kutainese	1	0.64%	84.71%
	Madurese	1	0.64%	85.35%

Demographic	N	Percentage	Cumulative
Malays	10	6.37%	91.72%
Minang	1	0.64%	92.36%
Chinese	12	7.64%	100%

Adaptation Procedure

Adaptation follows the guidelines of Beaton et al. (2000): there are six stages in the adaptation process: (1) Translation, (2) Synthesis, (3) Back-translation, (4) Expert judgment, (5) Pilot test, and (6) Psychometric property testing and report writing. The first part of translation is forward-translation by translating the dirty dozen instrument compiled by Jonason & Webster (2010) with English into Indonesian. Forward-translation was conducted with two experts. The first translator is an individual who has a bachelor's degree in English. Meanwhile, the second translator is an individual who works as a psychology lecturer familiar with the SD3 measurement tool. The second part of the translation is back-translation by translating the results of the first part of the translation in the form of Indonesian into English. The first translator was a lecturer who does not teach psychology. The second translator was a psychology lecturer. Data was then collected to test the psychometric properties of the adapted instrument.

Table 2. Sample Synthesis of T1 and T2 (T12) Items on the Dirty Dozen

No	Domain	Source	T1 Translation	T2 Translation	T12 (Synthesis)
04	Narcissism	I tend to want others to admire me	I like being praised/adored by others	I tend to want others to admire me.	I want others to admire me
02	Psychopathy	I tend to be callous or insensitive	I tend to be unfeeling or insensitive	I tend to be unfeeling or not sensitive.	I tend to be unfeeling or insensitive to others' feelings
10	Machiavellianism	I have used flattery to get my way	I tend to flatter others just to achieve my goals	I have used flattery to get what I want	I flatter others to achieve my goals

Note: T1 = Translator 1, T2 = Translator 2, T12 = Synthesis of T1 and T2

Expert Judgment

Of the 12 statement items assessed by expert judgment, there are 6 items that are appropriate if using the second back-translation result (T2). Statement numbers 1, 2, 3, 8, 10, and 11 are items that have T2 good enough to be used. For example, in item one with the original sentence, namely "i tend to want others to admire me" it is recommended to use the replacement sentence, namely "*Saya cenderung ingin orang lain mengagumi saya*". Some items assessed in the expert judgment only have one good enough translation, for example item number one. In addition to some statement items that were judged to have translational conformity (T2) with the original text, there were items that had conformity with the original text in the back-translation synthesis process. Some items, namely numbers 4, 6, 7, 10, and 11 had conformity when looking at the synthesis results with the meaning of the original text. For example, item number four with the original sentence "I tend to expect special favors from others". In the expert judgment process, the expert felt that the translated sentence that had meaning with the original text was the synthesized result (T12) with the sentence "*Saya cenderung mengharapkan bantuan istimewa dari orang lain*" (Table 2).
There were several items for which, during the expert judgment process, the expert suggested two translations that matched the meaning of the original text. The items are number 10 and 11. In item number 10, the original sentence reads, "I have used deceit or lied to get my way". In the expert judgment process,

the expert assessed that the results of T2 and T12 had meanings that were in accordance with the original text. T2 reads, "*Saya telah menggunakan tipu daya atau kebohongan*" and T12 reads, "*Saya menggunakan tipu daya atau kebohongan*". In this case, the researcher conducted a discussion to decide which translation result was more appropriate. In addition to the expert providing recommendations for translation results that are in accordance with the original text, the expert judgment process also produces new sentences that are not the same as the forward-translation and back-translation processes. However, the expert weighs the results of several translations, thus concluding a more suitable sentence. The writing of new sentences in the expert judgment process can be seen in items number 5, 9, and 12. Item number five has the original sentence, namely "I tend to lack remorse". From the translation results, the expert saw that there was no sentence that had a meaning close to the original text. Thus, the expert in the expert judgment process provided an alternative sentence, namely "*Saya cenderung kurang penyesalan*". Another example, namely in item number 12, the expert gave an alternative sentence, namely "*Saya memberikan sanjungan kepada orang lain untuk mencapai tujuan saya*".

Data Collection

Respondents filled out an online survey via google form. Respondents' willingness to participate in the study was asked at the beginning before filling in their personal data. Personal data contains gender, age, and ethnicity. After that, respondents filled out the survey page. The instrument used was the 12 items of the adapted Dirty Dozen instrument. Respondents were asked to read the statements and respond on 9 Likert points (1= Strongly disagree, 9 = Strongly Agree).

Data Analysis

Confirmatory factor analysis and reliability used R software with the *lavaan* package (Rosseel, 2012) to confirm the dark triad measurement model with three domains. Skewness and kurtosis were used to test the normality or distribution of the data. The acceptable values for data to be included in the normal distribution are skewness ±2 and kurtosis ±7 (Kim, 2013). The reliability of a good measuring instrument is as large as possible or close to the value of 1. The reliability of the measuring instrument can be categorized as quite consistent if the score is in the range of 0.7 to 0.8 (Nunnally & Bernstein, 1994). Confirmatory factor analysis with the criteria RMSEA < 0.10 (Hoyle, 2012), SRMR < 0.08 (Hu & Bentler, 1999), CFI and TLI > 0.90 (Hu & Bentler, 1999). Analysis of the effectiveness of response patterns using the Andrich Rating Scale Model (RSM) theory using WINSTEPS software (Linacre, 2023). According to Bond & Fox (2013) a comparison of each response category can be done to see the effectiveness of the response provided. The frequency of response categories to be tested is at least 10 responses and the average measurement increases constantly from the lowest response category to the highest response category. Infit and outfit mean square must be below 2.0. Values above this limit indicate that the response category does not contribute to the measurement of the latent variable but rather reduces the precision of the measurement. The threshold value that must be obtained by the response category is at least 1.4 logits greater than the previous response category. If the difference in value is more than 5 logits, it indicates that there is a difference in the object of measurement.

RESULTS AND DISCUSSION

Confirmatory Factor Analysis

Several analyses were conducted in factor analysis with a classical test theory approach to be applied to the dirty dozen instrument that has been adapted into Indonesian. Judging from the response patterns of the items in table 3, the items in the psychopathy and Machiavellianism domains were responded to in extremes. Both strongly disagree and strongly agree. This indicates the presence of response bias. It is known that some items have met the assumption of normality (P1, M4, N1, N2, N3 and N4), but some other items do not meet

the agreed normality assumptions (P2, P3, P4, M1, M2, and M3). If continued in confirmatory factor analysis will cause errors. An analysis was then conducted to test the internal consistency of the adapted instruments and found the respective Cronbach's alpha scores, namely the psychopathy domain of 0.37, Machiavellianism of 0.79, and narcissism of 0.92. The two domains of Machiavellianism and narcissism can be categorized as having good reliability.

Table 4. Goodness of Fit

No	Model	χ^2	RMSEA	SRMR	CFI	TLI
1	3 factors	131.523***	0.099	0.080	0.924	0.904
2	1 factor	429.028***	0.208	0.196	0.643	0.572
3	Second Order	141.154	0.103	0.089	0.916	0.895

) p < 0.05; **) p < 0.01; *) p < 0.001; χ2 = Chi-square; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index*

Confirmatory factor analysis was conducted to test the construct validity of the dirty dozen instrument model. The first model tried to test 3 factors (Figure 1). It was found that the first model met the fit criteria by setting constraints on item N2 fixed at 0.01. RMSEA is still within the agreed limit of <0.10 (Hoyle, 2012). However, the three fit indicators have met the fit criteria, SRMR < 0.08 (Hu & Bentler, 1999), CFI and TLI > 0.90 (Hu & Bentler, 1999). Testing the second model (Figure 2) was carried out because some articles argue that the dark triad is an interrelated construct, so researchers tried to prove by testing the dirty dozen as a unidimensional instrument. As can be seen in Table 4, all existing fit criteria cannot be met with a unidimensional model. The third test was conducted in second order form to test whether the dark triad consists of 3 sub-domains. The results showed that the model met several fit criteria. However, the narcissism domain was not able to measure the dark triad along with the other two domains. Based on the three findings above, the three domains (narcissism, psychopathy, and Machiavellianism) are domains that measure different constructs.

Figure 1. The first model

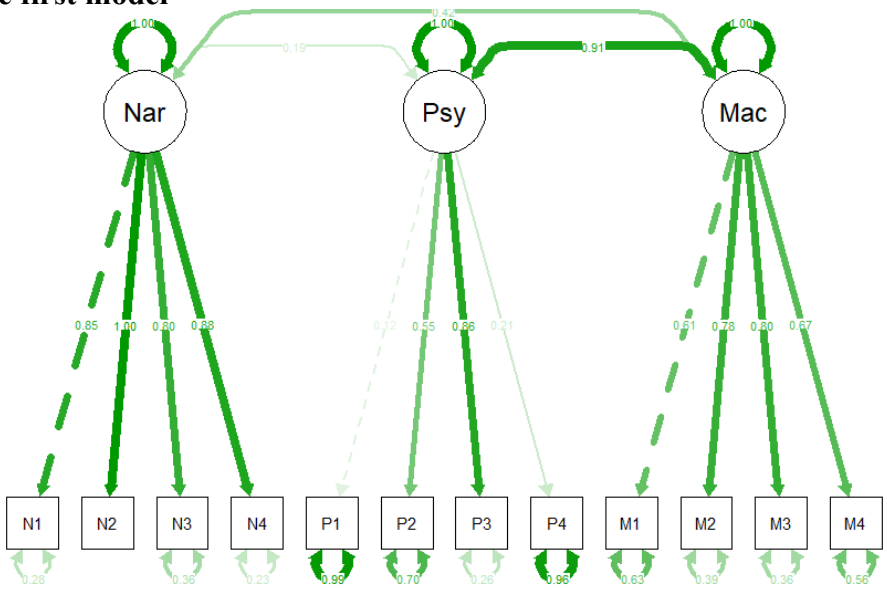
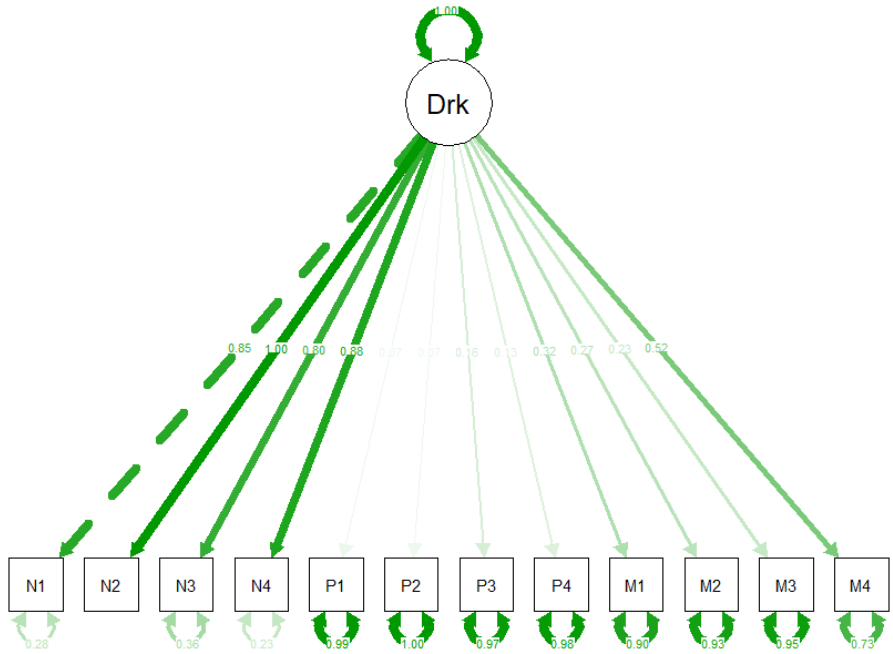


Figure 2. The second model



Comparative Analysis

Table 5. Mean (SD), Gender comparison, correlations

Domain	N=157	Laki-laki (N=61)	Perempuan (N=96)	t	1	2
1. Narcissism	19.60 (7.96)	19.16 (7.95)	19.88 (7.99)	0.545	-	-
2.Psychopathy	14.30 (4.61)	14.57 (4.56)	14.12 (4.66)	-0.607	0.190*	
3.Machiavellianism	11.38 (6.27)	11.23 (5.39)	11.47 (6.80)	0.232	0.441***	0.543***

) p < 0.05; **) p < 0.01; *) p < 0.001; a: komparasi jenis kelamin, df = 155*

Table 5 shows that the Machiavellianism domain has a fairly strong correlation with the psychopathy domain, the narcissism domain has a weak correlation with the psychopathy domain. This indicates an overlap between the Machiavellianism domain and the psychopathy domain. Behavioral characteristics in the Machiavellianism domain are equally indicative of characteristics from the psychopathy domain. Comparative analysis between the scores of men and women showed no difference in scores between men and women in the three domains.

Table 3. Item properties of the Indonesian version of Dirty Dozen

Items	Response scale (percent)									Descriptive				CITC	CID	Factor Loadings (λ)	Cronbach's α
	1	2	3	4	5	6	7	8	9	M	SD	Skewness	Kurtosis				
N1	6.37%	10.19%	10.19%	13.38%	8.92%	15.92%	22.93%	8.92%	3.18%	5.11	2.21	-1.42	-2.65	0.769	0.907	0.849	
N2	3.82%	12.74%	16.56%	9.55%	9.55%	17.83%	21.66%	7.64%	0.64%	4.90	2.09	-0.94	-3.19	0.964	0.843	0.999	0.918
N3	7.01%	14.01%	13.38%	10.19%	8.92%	13.38%	14.65%	14.65%	3.82%	4.95	2.39	-0.33	-3.32	0.724	0.926	0.802	
N4	7.01%	14.65%	16.56%	8.28%	10.19%	21.02%	14.01%	5.73%	2.55%	4.63	2.18	-0.02	-2.91	0.808	0.894	0.877	
P1	0.64%	8.28%	11.46%	12.10%	6.37%	9.55%	23.57%	16.56%	11.46%	5.90	2.24	-1.84	-2.90	0.017	0.519	0.120	
P2	23.57%	32.48%	24.20%	4.46%	3.18%	7.01%	3.18%	0.64%	1.27%	2.76	1.78	7.16	4.43	0.410	0.070	0.547	0.366
P3	47.13%	26.12%	13.38%	7.64%	1.27%	1.27%	1.27%	1.91%	0.00%	2.08	1.50	10.04	11.80	0.341	0.186	0.862	
P4	18.47%	24.84%	16.56%	7.64%	9.55%	10.83%	7.01%	0.64%	4.46%	3.55	2.24	3.84	-0.95	0.128	0.391	0.207	
M1	30.57%	26.12%	15.29%	10.83%	2.55%	4.46%	6.37%	3.82%	0.00%	2.87	2.02	5.77	0.88	0.529	0.775	0.610	
M2	42.04%	24.84%	16.56%	7.64%	1.27%	0.64%	1.91%	3.82%	1.27%	2.38	1.86	9.65	9.07	0.702	0.692	0.784	0.791
M3	47.13%	24.84%	15.29%	4.46%	2.55%	1.91%	1.27%	0.64%	1.91%	2.16	1.68	10.98	13.84	0.658	0.722	0.799	
M4	17.20%	19.75%	13.38%	8.92%	8.92%	14.01%	10.19%	3.82%	3.82%	3.98	2.38	2.00	-2.65	0.563	0.772	0.666	

CITC: Corrected item total correlation; CID: Cronbach's alpha if item deleted

Rating Scale Model Analysis

Table 6 shows the analysis using WINSTEPS (Linacre, 2023) to determine the effectiveness of response categories in response selection analysis based on Andrich Rating Scale Model (RSM) theory. The results show that the nine categories provided meet the first criterion of constant improvement. All responses meet the minimum response of 10. All responses show infit and outfit mean square values below 2.0, which means that the measurement error is still within reasonable limits. The threshold shows the effectiveness of the response options. The threshold value of the response category still does not meet the minimum limit of 1.4 logits.

Table 6. Response pattern on rating scale model

Category	Number of responses	Average measure	Infit mean square	Outfit mean square	Threshold	Threshold distance
1 (STS)	394	-1.19	0.89	0.93	None	-
2	375	-0.84	1.10	1.31	-0.95	0.95
3	287	-0.62	0.94	0.99	-0.46	0.49
4	165	-0.37	1.01	0.98	0.05	0.51
5	115	-0.23	0.98	0.98	0.05	0.00
6	185	0.00	0.82	0.81	-0.60	0.65
7	201	0.17	0.79	0.76	-0.04	0.56
8	108	0.28	1.10	1.18	0.84	0.88
9 (SS)	54	0.30	1.63	1.90	1.12	0.28

**) Units are log odd unit (logit)*

This study aims to adapt the Dirty Dozen (DD) scale into Bahasa Indonesia. DD is one of the instruments developed by Jonason & Webster (2010) to measure personality tendencies from a different perspective than other personality instruments, such as the big five inventory. The translation of the DD instrument conducted in this study aims to obtain a relatively standardized dark triad personality instrument that can be used for research purposes in Indonesia. The item analysis conducted showed that of the 12 items contained in the DD, there were two items from the psychopathy dimension, namely items P1 and P4 with low factor loadings (λ).

Item P1 reads "*Saya tidak mudah menyesal*" translated from "I tend to lack remorse". Based on the results of data collection, the participants' preference is to agree with this statement. Individuals with psychopathy tendencies will show no concern for the consequences of their actions. The tendency of participants to agree with this statement can be concluded as well as the occurrence of social desirability. The words "*Tidak mudah menyesal*" have not shown or described psychopathy tendencies, because it is feared that the word "*menyesal*" has a different meaning from "remorse". In the research of Tamura et al. (2015) obtained similar results. The psychopathy domain does not show good enough internal consistency due to the problem of equivalence of translation results into Japanese. It is necessary to adjust item P1 by providing a clearer context, such as "*Saya tidak menyesal ketika memukul orang tanpa sebab*", or using more equivalent sentences such as "*Saya tidak mudah merasa bersalah*".

Item P4 has a different back-translation from the original item which reads “I tend to be cynical”. When backtranslated, the item changed to “I doubt the good deeds of others” because the word “*sinis*” in the understanding of most Indonesians is different from “cynical”. Therefore, in the synthesis process, the researcher changed the item to “*Saya meragukan kebaikan orang lain*” which has a meaning more commensurate with the sound of the original item.

An interesting finding is the number of response options provided by the researcher. There were 9 response options provided. All respondents used the response options quite well. However, the results of the Andrich RSM analysis showed unsatisfactory results. The difference obtained between responses is too low. The low difference in threshold values for each category indicates confusion between the previous and next responses. For example, categories 4 and 5 are difficult for participants to understand because there is no clear enough difference. Adjustments need to be made in future research to create response categories that are suitable for the Dirty Dozen Indonesia instrument.

In the process of reviewing the theory, it was found that there were differences in the dark triad scores. In the research of Pineda et al. (2020) men tend to have higher scores than women. Especially in the psychopathy domain is the most contrasting thing in the dark triad. The tendency of psychopathy is a trait that is closely related to men. However, the results of the analysis of DD Indonesia show different results. There was no significant difference in scores between men and women. Both groups had similar characteristics when measured using the DD Indonesia instrument.

The results of the confirmatory factor analysis showed that the data had a fairly good fit with the proposed model. The 3-factor model remains better at measuring the dark triad. The correlation test also showed that there is a possibility of overlapping between the domains of psychopathy and Machiavellianism because these two domains are conceptually similar.

Despite the results of the confirmatory factor analysis of the DD Indonesia instrument model, two of the three domains have quite good measurement reliability, namely Machiavellianism (0.791) and Narcissism (0.918). This result is quite consistent and even better than the research of Kayiş et al. (2018) and Tamura et al. (2015). The domain that is quite difficult to reveal is the psychopathy domain. Ethical issues and desirability effects may have caused inconsistencies in the results.

In addition, researchers used two experts each in the forward-translation and back-translation processes. This is done so that the conclusions obtained from the translation results are objective. Researchers chose experts who have a background in psychology because it minimizes the possibility of failing to understand the concept of the theory used. In addition, the selection of experts who are unfamiliar with the science of psychology serves to minimize excessive confidence in the concept of psychological theory in experts who have a background in psychology. Therefore, the four experts in this study have good English, and of the four experts, two experts have a good understanding of the concept of psychological theory, and the other is unfamiliar with the concept of psychological theory.

In the expert judgment process, future research can add experts who are appropriate to the context. Expert judgment in some studies is also called cultural validation. This process aims to make the sentences on the adapted measuring instrument compatible with the realities in society through the selection of words made by experts according to their empirical experiences. Younan et al. (2019) conducted cultural validation with 7 experts who have different experiences. Experts have various training certificates, different places of residence, and various age ranges. Thus, the input or responses given to the items vary according to empirical experience, the use of language according to domicile, and generational differences from across ages.

Some of the limitations of this study, namely (1) The research respondents were still limited to meet the criteria for confirmatory factor analysis according to the suggestions of Muthén & Muthén (2002), (2) some adapted statements were still problematic because they were difficult to understand due to the abstract context, (3) the number of response categories that were too many (9 points) made it difficult for respondents to distinguish between categories, it was necessary to make adjustments again for future research, (4) other

measuring instruments were not used to test the criterion validity both convergent and discriminant against DD Indonesia. Thus, further research to produce a better Indonesian version of DD Indonesia can be carried out.

CONCLUSION

This study aimed to adapt and validate the Dirty Dozen (DD) instrument for measuring dark triad personality traits—narcissism, psychopathy, and Machiavellianism—in the Indonesian context. The findings revealed that the Indonesian version of DD demonstrated good reliability and construct validity for the narcissism and Machiavellianism domains, while the psychopathy domain showed lower reliability and translation challenges, particularly due to cultural nuances and abstract terminology. Notably, the results also indicated overlapping constructs between Machiavellianism and psychopathy, as well as no significant gender differences in the dark triad scores, contrasting with several international studies. The main research contribution lies in providing a standardized, efficient, and culturally relevant tool for assessing dark triad traits in Indonesian populations, while also identifying areas for refinement in the psychopathy dimension and response scale design, thus offering guidance for future instrument adaptation and cross-cultural psychometric research.

REFERENCES

Akhtar, H., & Sumintono, B. (2023). Rasch analysis of the International Personality Item Pool Big Five Markers Questionnaire: Is longer better? Primenjena Psihologija, 16(1), 3–28. https://doi.org/10.19090/pp.v16i1.2401

Alper, S., Bayrak, F., & Yilmaz, O. (2021). Inferring political and religious attitudes from composite faces perceived to be related to the dark triad personality traits. Personality and Individual Differences, 182, 111070. https://doi.org/10.1016/j.paid.2021.111070

Asih, S. R., & Lutfiyah. (2023). The Dark Triad Personality in Relation to Cyberbullying: The Role of Self-Esteem as a Mediator [Hubungan Kepribadian Dark Triad dan Cyberbullying: Peran Self-Esteem Sebagai Mediator]. ANIMA Indonesian Psychological Journal, 38(1), 038104. https://doi.org/10.24123/aipj.v38i1.4113

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. Spine, 25(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bond, T. G., & Fox, C. M. (2013). Applying the Rasch Model. Psychology Press. https://doi.org/10.4324/9781410614575

Campbell, W. K., Reeder, G. D., Sedikides, C., & Elliot, A. J. (2000). Narcissism and Comparative Self-Enhancement Strategies. Journal of Research in Personality, 34(3), 329–347. https://doi.org/10.1006/jrpe.2000.2282

Campbell, W. K., Rudich, E. A., & Sedikides, C. (2002). Narcissism, Self-Esteem, and the Positivity of Self-Views: Two Portraits of Self-Love. Personality and Social Psychology Bulletin, 28(3), 358–368. https://doi.org/10.1177/0146167202286007

Christie, R., & Geis, F. L. (1970). Studies in Machiavellianism. Academic Press.

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

Fehr, B., Samson, D., & Paulhus, D. L. (1992). The construct of Machiavellianism: Twenty years later. In Advances in personality assessment, Vol. 9. (pp. 77–116). Lawrence Erlbaum Associates, Inc.

Furnham, A., Richards, S., Rangel, L., & Jones, D. N. (2014). Measuring malevolence: Quantitative issues surrounding the Dark Triad of personality. Personality and Individual Differences, 67, 114–121. https://doi.org/10.1016/j.paid.2014.02.001

Gamache, D., Savard, C., & Maheux-Caron, V. (2018). French adaptation of the Short Dark Triad: Psychometric properties and a head-to-head comparison with the Dirty Dozen. Personality and Individual Differences, 122, 164–170. https://doi.org/10.1016/j.paid.2017.10.027

Hare, R. D. (1999). Without conscience: The disturbing world of the psychopaths among us. In Without conscience: The disturbing world of the psychopaths among us. The Guilford Press.

Harrel, W. A. (1980). RETALIATORY AGGRESSION BY HIGH AND LOW MACHIAVELLIANS AGAINST REMORSEFUL AND NON-REMORSEFUL WRONGDOERS. Social Behavior and Personality: An International Journal, 8(2), 217–220. https://doi.org/10.2224/sbp.1980.8.2.217

Hasanati, N., & Istiqomah, -. (2019). Validation of Dark Triad Personality Scale. Proceedings of the 4th ASEAN Conference on Psychology, Counselling, and Humanities (ACPCH 2018). https://doi.org/10.2991/acpch-18.2019.94

Hoyle, R. H. (2012). Handbook of structural equation modeling. In Handbook of structural equation modeling. The Guilford Press.

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

Jonason, P. K., & Sherman, R. A. (2020). Personality and the perception of situations: The Big Five and Dark Triad traits. Personality and Individual Differences, 163, 110081. https://doi.org/10.1016/j.paid.2020.110081

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., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3). Assessment, 21(1), 28–41. https://doi.org/10.1177/1073191113514105

Kaumbur, E. S., Wismanto, Y. B., & Hardjanta, G. (2018). Relationship between Emotional Intelligence and Religiosity With Dark Triad of Personality of Corruption Prisoner in Kedungpane Semarang prison. Jurnal Konseling Dan Pendidikan, 6(2). https://doi.org/10.29210/121700

Kayiş, A., Yılmaz, M., Satıcı, S., & Çapan, B. (2018). Dirty Dozen Scale: a study of adaptation to Turkish university students. Anatolian Journal of Psychiatry, 1. https://doi.org/10.5455/apd.288773

Kemper, C. J., Trapp, S., Kathmann, N., Samuel, D. B., & Ziegler, M. (2019). Short Versus Long Scales in Clinical Assessment: Exploring the Trade-Off Between Resources Saved and Psychometric Quality Lost Using Two Measures of Obsessive–Compulsive Symptoms. Assessment, 26(5), 767–782. https://doi.org/10.1177/1073191118810057

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restorative Dentistry & Endodontics, 38(1), 52. https://doi.org/10.5395/rde.2013.38.1.52

Linacre, J. M. (2023). Winsteps® Rasch measurement computer program (Version 5.6.0). Winsteps.com

Malesza, M., Ostaszewski, P., Büchner, S., & Kaczmarek, M. C. (2019). The Adaptation of the Short Dark Triad Personality Measure – Psychometric Properties of a German Sample. Current Psychology, 38(3), 855–864. https://doi.org/10.1007/s12144-017-9662-0

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
Miao, C., Humphrey, R. H., Qian, S., & Pollack, J. M. (2019). The relationship between emotional intelligence and the dark triad personality traits: A meta-analytic review. Journal of Research in Personality, 78, 189–197. https://doi.org/10.1016/j.jrp.2018.12.004
Muthén, L. K., & Muthén, B. O. (2002). How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. Structural Equation Modeling: A Multidisciplinary Journal, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8
Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). McGraw-Hill.
Nuzulia, S., & Why, F. Y. P. (2020). When the Dark Shines: The Role of Dark Personality Traits in Leadership Role Occupancy and Hiring Decisions in a Collectivistic Culture. Social Psychological and Personality Science, 11(8), 1089–1100. https://doi.org/10.1177/1948550619893956
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
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
Rauthmann, J. F., & Kolar, G. P. (2012). How "dark" are the Dark Triad traits? Examining the perceived darkness of narcissism, Machiavellianism, and psychopathy. Personality and Individual Differences, 53(7), 884–889. https://doi.org/10.1016/j.paid.2012.06.020
Rhodewalt, F., & Peterson, B. (2009). Narcissism. In Handbook of individual differences in social behavior. (pp. 547–560). The Guilford Press.
Rosseel, Y. (2012). lavaan : An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2). https://doi.org/10.18637/jss.v048.i02
Safaria, T., Nuqul, F. L., Purwandari, E., Ratnaningsih, I. Z., Khairania, M., Saputra, N. E., Rahmawati, E. I., Esita, Z., Nazriani, D., Miftahudin, M., & Mariati, L. I. (2020). The Role Of Dark Triad Personality On Cyberbullying: Is It Still A Problem? International Journal of Scientific & Technology Research, 9, 4256–4260. https://api.semanticscholar.org/CorpusID:222131044
Szabó, Z. P., Diller, S. J., Czibor, A., Restás, P., Jonas, E., & Frey, D. (2023). "One of these things is not like the others": The associations between dark triad personality traits, work attitudes, and work-related motivation. Personality and Individual Differences, 205, 112098. https://doi.org/10.1016/j.paid.2023.112098
Tajmirriyahi, M., Najafi, M., Hamidizadeh, K., Doerfler, S., & Ickes, W. (2021). The Dark Triad of personality and ideal romantic partner preferences in Iran. Personality and Individual Differences, 168, 110281. https://doi.org/10.1016/j.paid.2020.110281
Tamura, A., Oshio, A., Tanaka, K., Masui, K., & Jonason, P. K. (2015). Development, Reliability, and Validity of the Japanese Version of the Dark Triad Dirty Dozen (DTDD-J). The Japanese Journal of Personality, 24(1), 26–37. https://doi.org/10.2132/personality.24.26
Williams, K. M., & Paulhus, D. L. (2004). Factor structure of the Self-Report Psychopathy scale (SRP-II) in non-forensic samples. Personality and Individual Differences, 37(4), 765–778. https://doi.org/10.1016/j.paid.2003.11.004
Younan, L., Clinton, M., Fares, S., & Samaha, H. (2019). The translation and cultural adaptation validity of the Actual Scope of Practice Questionnaire. Eastern Mediterranean Health Journal, 25(3), 181–188. https://doi.org/10.26719/emhj.18.028
Zhuang, W.-L., Wu, S.-C., Wang, L.-H., & Huan, T.-C. (2022). Exploring the relationship between the Dark Triad personality traits and voice behaviour of hotel employees with the moderating effect of

intention to leave. International Journal of Hospitality Management, 107, 103294. https://doi.org/10.1016/j.ijhm.2022.103294