Jurnal Psikologi | E-ISSN: 2302-1098
2024, Vol. 23, No. 1, 208-220
https://doi.org/10.14710/jp.24.2.208-220

Quality test of IPIP-NEO Personality instrument items for Indonesian sample (Develop short version 90-items for Indonesian sample)

Abdurosid Nur Ali[1], Hanif Akhtar[1], Tulus Winarsunu[1], Silfiasari[1], Dian Caesaria Widyasari[2]
[1]Faculty of Psychology, University of Muhammadiyah Malang, Malang, Indonesia
[2]School of Psychology, University of Sheffield, South Yorkshire, England

ABSTRACT
Background: Increasing popularity of personality research on facet level, followed by instrument developments such as IPIP-NEO, that freely usable and modifiable measures. Research using the IPIP-NEO instrument is often conducted in Indonesia.
Purpose: This study aims to analyze the quality of items and develop a short version of the IPIP-NEO-300 for Indonesian samples that measures thirty facets within Big Five personality domains. The processes undertaken in this study include item analysis using the Rasch Model, the development of a shorter scale, and validation of this shortened version.
Method: The total subjects in this research is 759 people with an age range of 11-46 years. The sampling technique used for subject selection is convenience sampling, and data collection was carried out online using google forms.
Findings: These results indicate that IPIP-NEO 300 and IPIP-NEO 90 are psychometrically acceptable and practically convenient for measuring Big Five Personality on the facet level. IPIP-NEO short version 90 has a high correlation with IPIP-NEO 300, which means that IPIP-NEO short version can be used to measure personality in Indonesian samples.
Implication: The results of this study can enrich the psychometric literature in Indonesia and provide a basis for more specific and contextual personality research.

KEYWORDS
Big five, IPIP, NEO, rasch model

ARTICLE HISTORY
Received 11 February 2024
Revised 26 May 2024
Accepted 30 October 2024

Introduction

Personality is a psychological construct that is still developing in the world. Personality is a construct that has been highly demanded and extensively studied over the past decade. The Big Five personality theory is one of the popular personality theories used by researchers and psychologists to conduct assessments. McCrae (2010) explains in his Five Factor Theory (FFT) that personality is a system that lies between biological and socio-cultural conditions, including behavioural tendencies and character adaptations (habits, attitudes, roles, etc.). Research on the Big Five personality tends to only focus on domains and ignores specific traits of personality. A study by Pilarska (2018) explained that the Big Five Inventory (BFI) personality instrument, which uses the concept of Big Five Model (BFM), was used for practical purposes in describing variables and measuring self-concept and self-confidence. Consequently, the individual approach is weak in assessing personality.

Costa and McCrae (1985), using Big Five personality theory, created a measurement tool called the Neuroticism Extraversion Openness-Personality Inventory (NEO-PI). Initially, NEO-PI only had three domains; then an update was made by refining two additional factors, namely agreeableness and conscientiousness, forming NEO-PI-R (Costa & McCrae, 1992). The addition of new factors has led to further development of the Five Factor Model (FFM), which suggests that

CONTACT Abdurosid Nur Ali ✉abdurosidnurali@gmail.com ✉Jl. Raya Tlogomas No. 246 Malang, Jawa Timur, Indonesia

personality not only focuses on generalised domains but also represents a complex taxonomy. The introduction of the concept of facets in the NEO-PI-R (Costa & McCrae, 1992) is a distinct advantage of the NEO concept, specifically accelerating the use and inference of individual personality. A facet is a derivative or behavioural indicator of each domain, and it is highly interrelated to be grouped into a domain. These facets better explain the general Big Five domains used in NEO-PI-R, with six facets each, totalling 30 facets. Costa and Mccrae (1995) developed the concept of facets with the aim of creating a clearer hierarchy to explain the formation of a domain. Therefore, the Five Factor Model (FFM) can comprehensively explain all human traits through a complex concept of facets rather than in more general domains. This is an advantage of the NEO concept over other personality concepts.

The development of the Big Five Personality Theory led to the development of several instruments, such as the Big Five Inventory (John & Srivastava, 1999), Trait Descriptive Adjective-TDA (Goldberg, 1999), and NEO-PI (McCrae & Costa, 1987). In general, the NEO has the best validity in measuring the Big Five personality theory. The BFI is commonly used in research due to its concise items and efficient administration, which provides better context. The TDA utilised a single word of 100 adjective items to assess personality, which is simpler but lacks detail in describing personality compared to the other two instruments. All three instruments have also been widely adapted to various languages, especially in Europe, and demonstrated high reliability in measuring the five domains of personality. Despite being widely used, the high reliability rooted in a large number of items accentuates the need for shorter and more effective alternatives to measure personality. Another major obstacle to using the NEO-PI-R is purchasing the license, which effectively increases the exclusivity of its usage and hinders the development of science.

The International Personality Item Pool (IPIP) is one of the platforms that helps the development of personality measurement tools in the world, and for free. The available items on the IPIP website measure the same constructs as the original instrument, such as the IPIP-NEO, which also measures the same constructs as the NEO-PI-R (Goldberg et al., 2006) and it is currently being adapted to various languages. IPIP-NEO is a free or open-source version of NEO-PI-R developed by Costa and McCrae (1995). Goldberg et al., (2006) explained that the statement items on the IPIP-NEO instrument have been selected to represent the original instrument. It is exemplified that the high correlation between items in NEO-PI-R (Costa & McCrae, 1992) and IPIP is .94. According to Ashton (Goldberg et al., 2006), a high correlation between the original instrument and the constructed one can indicate alignment within the construct.

The Indonesian IPIP-NEO test is an adaptation of the English version of the IPIP NEO test. This test is an open-source instrument that measures construct equivalent to the NEO-PI-R, which is available on the IPIP website (Goldberg et al., 2006). The IPIP-NEO provides a 300-item long version measuring all 30 facets from five domains of the Big Five personality, namely, neuroticism, extraversion, openness, agreeableness, and conscientiousness. Neuroticism is a personality filled with negative emotions such as anger and anxiety, extraversion is a personality that is enthusiastic, positive, and likes activities related to other people, openness is a personality that is open to new things, agreeableness is a personality that is friendly and generous, while conscientiousness is a personality that likes things that are orderly and organized (Sangwan, 2023).

Goldberg (1999) published the results of his research in compiling the IPIP-NEO instrument, which has the same construct and focuses on the 30 facets that are the same as those measured in the NEO-PI-R instrument. Table 1 displays the domains and facets of IPIP-NEO.

Table 1.
Domains and Facets of IPIP-NEO

Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Anxiety	Friendliness*	Imagination*	Trust	Self-efficacy*
Anger*	Gregariousness	Artistic interest*	Morality*	Orderliness*
Depression	Assertiveness	Emotionality*	Altruism	Dutifulness
Self-Consciousness	Activity-level*	Adventurousness*	Cooperation*	Achievement-striving
Immoderation*	Excitement-seeking	Intellect*	Modesty	Self-discipline
Vulnerability	Cheerfulness*	Liberalism*	Sympathy*	Cautiousness

(*) Different facet naming with NEO-PI-R

There are 16 facet names, and they differ in both instruments, but it does not affect the meaning of the NEO constructs measured in NEO-PI-R. Despite better psychometric properties, using the original long version is inefficient due to the unavoidable time-consuming, high cost, complicated administration, and boredom in filling out the instrument by respondents. According to Herzberg & Brähler (2006), researchers tend to aim for shorter, more efficient, and reliable research instruments without compromising their construct validity, including for the IPIP-NEO. One study uses the IPIP-NEO instrument, as seen in Vedel et al., (2019) on Danish subjects. The results show that the use of the IPIP-NEO with 120 items met the validity and reliability criteria and can be used in Danish samples. Another study by Kajonius and Johnson (2019) with US samples shows that IPIP-NEO is structurally robust for research and clinical practice in personality assessment. In Indonesia, a study on IPIP-NEO was conducted by Tarigan (2024) who reported the validity/reliability results of IPIP-NEO, which showed that 120 items could explain the construct of the Big Five Personality.

In Indonesia, there are some studies on personality and its relationship to other variables. Research by Yulisa et al., (2023) explains that there is a relationship between Big Five personality and school adjustment to aggressiveness towards teachers, where only agreeableness personality is related to aggressiveness towards teachers. Another research by Simanullang (2021) explains that there is an influence of the Big Five personality model personality type on the performance of state civil servants. Murhadi et al., (2024) conducted research on Indonesian investors during the COVID-19 pandemic. The results showed that investors who are more extroverted, open to new experiences, and more conscientious/organized tend to be more tolerant of financial risk and make bolder investment decisions during volatile market conditions such as the pandemic.

Nevertheless, the use of the IPIP-NEO instrument in research in Indonesia remains minimal, and there are no instruments that comprehensively measure the 30 facets of the Big Five in Bahasa Indonesia. The most used instruments, such as the Ten-Item Personality Inventory (TIPI), Big Five Inventory (BFI), IPIP-BFM, etc., are copyrighted and are not free to use. In addition, unclear psychometric properties of these personality instruments also cause problematic use in Indonesian samples. Given that the IPIP-NEO is widely used, free, and measures 30 facets from five domains of the Big Five personality theory, it offers more benefits to be adapted to Bahasa Indonesia.

This study aims to test the quality of the IPIP-NEO adaptation items with the Rasch Model approach and develop a short version personality measurement tool. The short version will

comprise as many as 120 items from IPIP-NEO by maintaining 30 facets that have been adopted by the Laboratorium Psikologi (2020). We hypothesized that the short version of IPIP-NEO has adequate psychometric quality (validity and reliability) to be used in Indonesian samples. This study will provide evidence for the 90-item IPIP-NEO testing on Indonesian samples, and cater to the need for a concise and complete personality instrument with clear psychometric properties. This instrument will be beneficial in various fields, such as clinical personality assessment, corporate employee recruitment, learning in an academic context, and facilitating researchers who wish to utilise personality instruments as study variables.

Method

This research method is a descriptive quantitative study aimed at developing a short version of 90 items of the IPIP-NEO Indonesian version and testing the psychometric properties, including item quality assessment and instrument reliability. Data collection was carried out online using google forms, which was then distributed through lecture classes and social media in Indonesia. Respondents are given freedom in the use of data collected by the researcher, such as requesting deletion and confidentiality by contacting the researcher after they submitted responses. Subjects were also asked to complete an informed consent form on Google Form, which sought their willingness to participate in the research.

The target subjects of this study were male and female, aged 11 years or older, and Indonesian citizens residing in various regions in Indonesia (n = 759; 67.5% female; 32.5% male). The sampling technique used for subject selection is convenience sampling, where the subjects in this study are subjects who are accessible to researchers and are willing to be involved in the study (Golzar et al., 2022). Considering the recommended subjects in the analysis using the Rasch Model according to Azizan et al., (2020) is 500 subjects to achieve satisfactory maximum likelihood estimation (MLE). This study included a total of 759 participants.

Table 2.

Characteristics of Research Participants (N=759)

Characteristics	N	Percentage (%)
Gender		
Male	247	32.5
Female	512	67.5
Age		
11-20	608	80.1
21-30	148	19.5
31-40	1	.1
41-50	0	-
51-60	2	.3
Hometown		
Bali	5	.7
Jawa	660	87
Kalimantan	48	6.3
Lombok	3	.4
Maluku	5	.7
Nusa Tenggara	7	.9
Papua	2	.3
Sulawesi	10	1.3
Sumatera	19	2.5

The instrument used in this study is the IPIP-NEO compiled by Goldberg (1999) with 300 statements that have been adapted by the Laboratorium Psikologi (2020) spreaded across five domains: neuroticism (.94), extraversion (.92), openness (.79), agreeableness (.84), and conscientiousness (.93). Examples of neuroticism items: 'Often feels sad'; extraversion: "Feels comfortable among many people"; openness: 'Has a vivid imagination'; agreeableness: 'Likes to help others'; conscientiousness: 'Completes tasks thoroughly'. The instrument is in Likert form, where favourable items have a score of 1, meaning strongly disagree (STS) to 5, meaning strongly agree (SS), and on unfavourable items, the opposite applies.

The analysis of item quality for the IPIP-NEO was conducted using the Rasch model. To create the short version of the instrument, three items were selected from each facet. Each item was evaluated based on several criteria, including the level of item fit, which is measured by the average values of infit and outfit (Khotimah et al., 2024). Additionally, item difficulty and item discrimination power were assessed for each facet during the item selection. To evaluate the reliability of the 90-item IPIP-NEO, the Cronbach's alpha coefficient was calculated. We also measure the separation of person and item. Finally, a Pearson correlation test was performed to examine the overlap between the long version (300 items) and the short version (90 items) IPIP-NEO (Dufera et al., 2023).

Result

Tests were conducted on various items, specifically on the item fit to the model, point-measure correlation, and item discrimination for each facet, totalling 10 items (see Table 3). Items with an outfit mean square value outside the range of .5 and 1.5 indicate item misfits or items that are not suitable for measurement. Additionally, items with a point-measure correlation or discrimination power value between .4 and .85 are effective at discriminating subjects. Both parameters can identify items that are misfitting or inappropriate for measurement.

Table 3.
Item Analysis Summary IPIP-NEO-300

Domain/faset	*Measure (Logit)*		*Outfit Mnsq*		*Pt. Measure Correlation*		Item *misfit*
	Min	*Max*	*Min*	*Max*	*Min*	*Max*	
Neuroticism							
Anger	-.34	1.12	.71	1.48	.48	.78	0
Anxiety	-.92	1.32	.8	1.72	.27	.72	2
Depression	-1.01	.9	.74	.74	.33	.71	2
Immoderation	-1.49	.99	.83	1.15	.23	.6	4
Self-Consciousness	-.91	.8	.81	1.39	.34	.64	3
Vulnerability	-.82	.75	.65	1.33	.45	.63	0
Extraversion							
Activity-Level	-.78	1.23	.77	1.4	.16	.54	3
Assertiveness	-.55	.57	.7	1.21	.43	.59	0
Cheerfulness	-.42	1.18	.72	1.65	.29	.62	2
Excitement Seeking	-.69	1.12	.71	1.92	.34	.6	2
Friendliness	-1.03	.64	.69	1.58	.34	.7	1
Gregariousness	-.75	.77	-.75	1.64	.34	.34	1

Domain/faset	*Measure (Logit)*		*Outfit Mnsq*		*Pt. Measure Correlation*		Item *misfit*
	Min	*Max*	*Min*	*Max*	*Min*	*Max*	
Openness							
Adventurousness	-1.25	1.5	.79	1.18	.32	.63	1
Artistic Interest	-1.35	.64	.75	1.5	.34	.63	2
Emotionality	-.99	.81	.72	1.48	.29	.52	4
Imagination	-.63	.62	.86	1.22	.38	.57	1
Intellect	-1.13	.57	.76	1.19	.32	.58	1
Liberalism	-1.52	1.28	.79	1.36	.2	.47	6
Agreeableness							
Altruism	-.77	.97	.69	1.27	.42	.6	0
Cooperation	-.83	.56	.65	1.26	.14	.66	3
Modesty	-.84	.69	.77	1.16	.38	.6	1
Morality	-.75	1.52	.71	2.1	.2	.68	3
Sympathy	-.9	1.31	.82	1.23	.26	.47	3
Trust	-1.07	1.59	.71	1.41	.37	.66	1
Conscientiousness							
Achievement Striving	-.64	1.43	.66	1.6	.11	.61	2
Cautiousness	-.61	.96	.65	1.65	.28	.7	2
Dutifulness	-.55	.92	.79	1.28	.41	.61	0
Orderliness	-.52	1.01	.73	1.31	.44	.64	0
Self Discipline	-.82	1.51	.78	1.49	.48	.73	0
Self Efficacy	-.66	.87	.72	1.46	.36	.63	1

Item selection for the IPIP-NEO 90 scale involves selecting three items from each facet that demonstrate good suitability, discrimination power, and appropriate difficulty. The IPIP-NEO 90 will be tested for each factor individually, for example, the neuroticism factor consists of 18 items, with three items selected from each of its facets.

The analysis stage included testing the dimensionality of each domain within the 60-item IPIP-NEO. The results of the dimensionality analysis in Table 5 revealed that the items explained an average of 32% of the variance within each domain. Additionally, a dimensionality analysis was carried out on the IPIP-NEO-90, which consists of 18 items per domain, showing an average of 34.7% variance explained by the IPIP-NEO 90. Instruments demonstrating a dimensionality above 30% are generally considered acceptable (Tennant & Küçükdeveci, 2023).

Table 4.

Dimensionalitas IPIP-NEO

Domain	IPIP-NEO 300		IPIP-NEO 90	
	Variance Explained	*Variance Unexplained*	*Variance Explained*	*Variance Unexplained*
Neuroticism	32.2%	67.8%	34.7%	65.3%
Extraversion	32.4%	67.6%	35.1%	64.9%
Openness	33.8%	66.2%	41.4%	58.6%
Agreeableness	30.6%	69.4%	30.7%	69.3%
Conscientiousness	31.1%	68.9%	31.6%	68.4%

The results of the reliability analysis (see Table 5) generated by WINSTEPS calculate person and item reliability measures. Person reliability indicates the consistency of the subject's ability being measured, while item reliability reveals the consistency of the measurement items themselves. The reliability of the instrument is indicated by the Cronbach's alpha value.

Table 5.

Reliability IPIP-NEO 300 and 90

Domain	*Average Logit Measure*		*Separation*		*Reliability*		α	
	300	**90**	**300**	**90**	**300**	**90**	**300**	**90**
Neuroticism								
Person	.12 (.50)	.07 (.62)	3.3	2.02	.92	.8	.93	.82
Item	.00 (.49)	.00 (.44)	12.98	11.55	.99	.99		
Extraversion								
Person	.28 (.43)	.29 (.55)	2.8	1.71	.89	.75	.9	.77
Item	.00 (.54)	.00 (.53)	14.05	13.63	.99	.99		
Openness								
Person	.34 (.27)	.16 (.33)	1.65	.76	.73	.37	.77	.44
Item	.00 (.66)	.00 (.76)	16.96	19.54	1	1		
Agreeableness								
Person	.51 (.33)	.66 (.49)	2	1.36	.8	.65	.83	.7
Item	.00 (.58)	.00 (.54)	14.86	13.08	1	.99		
Conscientiousness								
Person	.58 (.58)	.58 (.68)	3.19	1.94	.91	.79	.93	.81
Item	.00 (.49)	.00 (.41)	11.91	10.04	.99	.99		

Note: Score average logit measure M (SD); 300 = IPIP-NEO-300; 90 = IPIP-NEO-90

Correlation tests were conducted using the total scores from each domain of the IPIP-NEO 300 and IPIP-NEO 90 scales. The results of these analyses are shown in Table 6. The correlation within the same domain ranged from .816 to .949, indicating a strong correlation between IPIP-NEO 300 and IPIP-NEO 90. Both scales align well in their measurement and representation capabilities.

Tabel 6.

Correlation of Each Domain IPIP-NEO 300 and 90

	N_90	**E_90**	**O_90**	**A_90**	**C_90**
N_300	**.947***	-.419*	.147*	-.063	-.483*
E_300	-.384*	**.949***	.165*	.053	.325*
O_300	-.061	.312*	**.816***	.221*	.260*
A_300	-.013	.015	.056	**.911***	.340*
C_300	-.487*	.328*	.069	.399*	**.940***

Note: N=Neuroticism; E=Extraversion; O=Openness; A=Agreeableness; C=Conscientiousness; Significant correlation at the level 0.001 (2-tailed)

Discussion

The aim of this study was to examine the item quality of the IPIP-NEO instrument using the Rasch Model approach. This analysis allows researchers to further discuss the psychometric properties of IPIP-NEO from a modern test theory perspective. A series of analyses were conducted, demonstrating that the IPIP-NEO 300 has satisfactory psychometric properties. The

reliability values obtained were adequate, both using the internal consistency approach and the Rasch Model.

The reliability coefficient of this study using the Rasch Model obtained a value that was similar to the classical test theory approach. In research conducted by the Laboratorium Psikologi (2020), the Cronbach's alpha reliability ranged from .785 to .938. The reliability of the IPIP-NEO 300 through the Rasch Model approach can be seen through the person reliability, which ranges from .73 to .92. Meanwhile, IPIP-NEO-90 has a reliability value ranging from .44 to .82. The difference is minimal due to the various ways of calculation employed. The Cronbach's alpha approach uses raw scores, while Rasch uses logit scores or scores that have been converted into intervals. According to Tunç (2023) analysis with the Rasch model for adaptation of measuring instruments is more recommended because it has many benefits, such as measuring unidimensionality, respondent and item reliability (person/item separation), and examining DIF (Differential Item Functioning) in a cross-cultural context.

Furthermore, within the context of the Rasch Model, evaluation of the unidimensionality and quality of an instrument is generally assessed through the proportion of variance explained by measures. A value above 30% is often considered to indicate a primary dimension that is sufficiently robust to represent the construct being measured (Walpuski & Celik, 2024). However, recent studies emphasize that this threshold is not an absolute criterion, but rather a preliminary indicator that needs to be supplemented with residual and item fit analyses (Davis & Boone, 2021). The short version of the IPIP-NEO was concluded to have a variance explained by measures value above 30%, indicating that all dimensions adequately explain the Big Five personality traits.

The procedure for preparing the IPIP-NEO 90 short instrument in this study did not closely replicate the research conducted by Maples-Keller et al., (2019), which performed item fit and item discrimination tests to identify inappropriate items in the measurement. The different samples in this study show various results. This finding is supported by research by Chen et al., (2016), revealing that personality domains in samples that have different cultures produce opposite findings. In the Openness domain, unsatisfactory results were obtained; the reliability obtained was .77 on IPIP-NEO 300 and .44 on IPIP-NEO 90. This discrepancy may arise from the fact that the items on the IPIP-NEO were developed within a Western setting, specifically America and Europe, which may not resonate with the communal values found in Indonesian samples. For instance, one item related to the liberalism facet states, 'Tend to vote for liberal political candidates'. This item can create problems for Indonesian samples, as agreeing with this statement may be difficult due to the conservative, collectivist political culture prevalent in Indonesia. Expressing overly liberal opinions is often met with negative sentiment in Indonesia, especially when voicing political views.

Moreover, a low discrimination score below .4 indicates that the item cannot differentiate between the subject groups, suggesting a lack of variation in the individuals tested in the data collection process. The low discrimination shown in item 199 in the immoderation facet, which states 'enjoying food', indicates that the respondents are likely to unanimously agree with it, resulting in no group variation. Item 229 on the activity level facet, which states 'Not in a hurry' with a point measure correlation value of .16, indicates that the item does not accurately measure the facet due to its ambiguous context, making it difficult for respondents to understand. Consequently, both items mentioned are concluded not to fit or misfit, as they cannot precisely measure the facet in Indonesian samples.

According to research by Laajaj et al., (2019), it was found that, generally, the statements on the personality instrument failed to measure the intended trait and demonstrated low validity. This issue stems from these personality instruments have typically been developed in an educated

population (up to university level) with a high per capita income. In contrast, Indonesia is classified as a non-WEIRD (western, educated, industrialised, rich, and democratic) country, which may result in inconsistencies between the personality assessment and the actual realities of the population (Mo˜ttus & Allerhand, 2018).

This study provided more information about the psychometric properties of IPIP-NEO by employing the classical test theory and the modern test theory approaches. Although the IPIP-NEO 300 items generally exhibit satisfactory psychometric properties for use in Indonesian samples, some items show unsatisfactory parameters, including misfit items or low point measure correlation values. These problematic items need to be addressed first, particularly the adaptability of the statements appropriate to the Indonesian context (Verpalen et al., 2018).

The domain of the IPIP-NEO-300 instrument is highly correlated with the shorter 90-item version, indicating good alignment between instruments to measure the same psychological constructs. For instance, in the extraversion domain, shown with a value of .949, it indicates that IPIP-NEO-90 can represent IPIP-NEO-300 well. In the neuroticism domain, the correlation value of IPIP-NEO-90 and IPIP-NEO-300 is 0.947, which shows that IPIP-NEO-90 can represent IPIP-NEO-300 well. Likewise, other domains, namely openness, agreeableness, and conscientiousness, have high correlation values, indicating that all IPIP-NEO-90 domains can explain the original version well. The IPIP-NEO-300 has slightly better reliability and item separation than the IPIP-NEO-90. Caution should be exercised when using the shorter version, even though the smaller number of items can adequately represent the long version. Moreover, this study also produced the IPIP-NEO-90, which fairly succinct Big Five personality instrument with good validity and reliability in Bahasa Indonesia. Providing an instrument that is accompanied by robust and satisfactory psychometric properties is essential for a study in this field. It is anticipated that research involving personality can grow rapidly and not be restricted.

Several recent studies in Indonesia on personality require a measuring instrument that can accurately measure personality. The IPIP-NEO-90 can be an alternative in personality measurement. The IPIP-NEO-90 consists of 90 items across five personality domains: neuroticism, extroversion, openness, agreeableness, and conscientiousness. Neuroticism refers to an individual's tendency to experience negative emotions such as anxiety, emotional instability, and vulnerability to stress (Nugroho & Huwae, 2025). In contrast, extroversion describes individuals who are energetic, sociable, and seek social stimulation, while individuals who score low on this dimension tend to be more reserved and reflective openness relates to intellectual curiosity, imagination, creativity, and openness to new ideas and experiences (Martinez et al., 2025). This trait is related to the ability to think divergently and a tendency to appreciate art, nontraditional values, and the exploration of complex ideas.

Meanwhile, agreeableness reflects prosocial traits such as empathy, warmth, willingness to cooperate, and a tendency to avoid interpersonal conflict (Ismaniar & Uyun, 2023). Individuals with high levels of agreeableness tend to be altruistic and maintain social harmony, while low scores are associated with competitive or antagonistic attitudes. Finally, conscientiousness is related to discipline, perseverance, and self-regulation, which are consistently linked to better academic and work performance (Furnham & Cheng, 2024).

IPIP-NEO has been adapted to various countries into a short version. Ritter et al., (2024) adapted the IPIP-NEO in Portuguese culture, resulting in 16 personality scales for work-related personality traits in Portuguese and professional contexts. Cupani and Lorenzo-Seva (2016) adapting the IPIP-NEO to Argentine culture, the findings suggest that the IPIP-Revised is a valid option for measuring personality traits in the Argentine population. This demonstrates that

adaptations of the short version of the IPIP-NEO in various countries can explain the Big Five personality dimensions just as well as the original version.

The short version of the IPIP-NEO-90 adaptation offers significant advantages for use in various contexts in Indonesia. This version offers an efficient alternative to the long 300-item form, maintaining the comprehensive structure of the Big Five Personality Scale while requiring a shorter administration time. This makes the IPIP-NEO-90 ideal for use in educational settings, such as student character assessments or personality psychology research in schools and universities. One example is research conducted by Erik (2020) which showed that agreeableness influences entrepreneurial intentions among students in Jakarta. The IPIP-NEO-90 can also be used in organizational settings, such as for new employee recruitment, performance evaluation, and human resource development. One example of Big Five personality research in an organizational setting is by Nasyroh and Wikansari (2017) which showed that all Big Five personality dimensions are related to employee performance within a company.

In clinical settings, the Big Five personality can be used for psychological assessment and intervention, as was done by Bucher et al., (2019) who showed that individuals with an openness personality are more receptive to experience-based therapy approaches such as art therapy or narrative therapy. Furthermore, the short version is also suitable for large population research or national surveys where time efficiency is a key consideration. In the context of Indonesian culture, with its diverse literacy levels and social backgrounds, this short form minimizes respondent fatigue and increases response accuracy without compromising the instrument's reliability and validity. Thus, the IPIP-NEO-90 can be an adaptive, practical, and relevant personality measurement tool to support psychology research and practice in Indonesia.

The limitation of this study is that the quality of the items from the adapted IPIP-NEO instrument is not fully representative. The research subjects were predominantly from the Java region, which means they do not adequately reflect the diversity of Indonesian samples. Some items also had poor parameter values, indicating that the respondents may not have fully understood the statements. An improvement is possible for future studies, particularly addressing the IPIP-NEO-90 domains with unsatisfactory reliability values. Increasing sample variability in subsequent research may offer more consistent and robust evidence for measuring NEO constructs in Indonesian samples. This study implied that the short version of the IPIP-NEO-90 can help psychology practitioners measure personality effectively. Since the short version has fewer items than the original, the IPIP NEO is particularly well-suited for large-scale applications, such as recruitment and psychological assessment.

Conclusion

This study concludes that the IPIP-NEO 300 has good psychometric properties. The IPIP-NEO 300 has satisfactory validity and reliability. IPIP-NEO 90 has a high correlation value with the IPIP-NEO-300, indicating the feasibility of IPIP-NEO 90 being used in the measurement of Big Five personality.

For future researchers studying the Big Five Personality, particularly using the IPIP-NEO instrument, it is important to improve the adaptation of some items in the instrument to better capture the Indonesian context. Under certain conditions, IPIP-NEO-90 can be used, but caution is advised due to its low reliability.

Credit/Acknowledgement

The researcher would like to express sincere gratitude to all participants who kindly contributed their time and effort to take part in this study. Their willingness to share experiences and insights has been invaluable to the completion of this research. The researcher would also like to extend deepest appreciation to the Psychology Laboratory University of Muhammadiyah Malang for the continuous support, facilities, and assistance provided throughout the research process. Such support made it possible for this study to run smoothly and successfully.

References

Azizan, N. H., Mahmud, Z., & Rambli, A. (2020). Rasch rating scale item estimates using maximum likelihood approach: Effects of sample size on the accuracy and bias of the estimates. *International Journal of Advanced Science and Technology*, *29*(4), 2526–2531.

Bucher, M. A., Suzuki, T., & Samuel, D. B. (2019). A meta-analytic review of personality traits and their associations with mental health treatment outcomes. *Clinical Psychology Review*, *70*, 51–63. https://doi.org/10.1016/j.cpr.2019.04.002

Chen, B.-Y., Wan, L., & Fang, Y. (2016). Interpersonal self-support traits predicting interpersonal trust among undergraduate students: Beyond the effects of big five personality. *Psychological Science*, *39*(6), 1441–1447.

Costa, P. T., & McCrae, R. C. (1992). *Revised NEO personality inventory (NEO PI-R™) and NEO five-factor inventory (NEO-FFI): Professional manual*. Psychological Assessment Press.

Costa, P. T., & McCrae, R. R. (1985). *The NEO personality inventory manual*. Psychological Assessment Resources.

Costa, P. T., & Mccrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment*, *64*(1), 21–50. https://doi.org/doi.org/10.1207/s15327752jpa6401_2

Cupani, M., & Lorenzo-Seva, U. (2016). The development of an alternative IPIP inventory measuring the big-five factor markers in an Argentine sample. *Personality and Individual Differences*, *91*, 40–46.

Davis, D. R., & Boone, W. (2021). Using Rasch analysis to evaluate the psychometric functioning of the other-directed, lighthearted, intellectual, and whimsical (OLIW) adult playfulness scale. *International Journal of Educational Research Open*, *2*, 1–12. https://doi.org/10.1016/j.ijedro.2021.100054

Dufera, A. G., Liu, T., & Xu, J. (2023). Regression models of Pearson correlation coefficient. *Statistical Theory and Related Fields*, *7*(2), 97–106. https://doi.org/10.1080/24754269.2023.2164970

Erik, E. (2020). Pengaruh kepribadian lima besar terhadap intensi berwirausaha pada mahasiswa psikologi Universitas Negeri Jakarta. *JPPP - Jurnal Penelitian dan Pengukuran Psikologi*, *9*(2), 96–101. https://doi.org/10.21009/JPPP.092.07

Furnham, A., & Cheng, H. (2024). Correlates of conscientiousness: Findings from the millennium cohort study. *The Journal of Genetic Psychology*, *185*(2), 114–123. https://doi.org/10.1080/00221325.2023.2279143

Goldberg, L. R. (1992). Development of markers for the big five factor structure. In *Psychological Assessment* (Vol. 4, Issue 1, pp. 26–42).

Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. *8th, European Conference on Personality*, *7*, 7–28.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Golzar, J., Noor, S., & Tajik, O. (2022). Convenience sampling. *International Journal of Education Language Studies*, *1*(2), 72–77. https://doi.org/10.22034/ijels.2022.162981

Herzberg, P. Y., & Brähler, E. (2006). Assessing the big-five personality domains via short forms a cautionary note and a proposal. *European Journal of Psychological Assessment*, *22*(3), 139–148. https://doi.org/10.1027/1015-5759.22.3.139

Ismaniar, I., & Uyun, M. (2023). Extraversion, agreeableness and gratitude. *Psikostudia: Jurnal Psikologi*, *12*(1), 132–137.

John, Oliver. P., & Srivastava, Sanjay. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research (2nd Ed.)*.

Khotimah, K., Hadi, T. S., & Lestari, I. (2024). Application of the rasch model in research publications: A bibliometric analysis. *Plusminus: Jurnal Pendidikan Matematika*, *4*(2), 229–240. https://doi.org/10.31980/plusminus.v4i2.1466

Laajaj, R., Macours, K., Hernandez, D. A. P., Arias, O., Gosling, S. D., Potter, J., Rubio-Codina, M., & Vakis, R. (2019). Challenges to capture the big five personality traits in non-WEIRD populations. *Science Advances*, *5*(7). https://doi.org/10.1126/sciadv.aaw5226

Laboratorium Psikologi. (2020). *Manual IPIP-NEO Indonesia*. Laboratorium Psikologi Universitas Muhammadiyah Malang.

Maples-Keller, J. L., Williamson, R. L., Sleep, C. E., Carter, N. T., Campbell, W. K., & Miller, J. D. (2019a). Using item response theory to develop a 60-item representation of the NEO PI–R using the international personality item pool: Development of the IPIP–NEO–60. *Journal of Personality Assessment*, *101*(1), 4–15. https://doi.org/10.1080/00223891.2017.1381968

Martinez, R. L., Regan, A., Gollwitzer, P. M., Oettingen, G., & Lyubomirsky, S. (2025). Increasing extraversion via intervention: Lay insights, person-activity fit, and implications for well-being and persistence. *Social and Personality Psychology Compass*, *19*(2), 1–12. https://doi.org/10.1111/spc3.70043

McCrae, R. R. (2010). The place of the FFM in personality psychology. *Psychological Inquiry*, *21*(1), 57–64. https://doi.org/10.1080/10478401003648773

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

Mo˜ttus, R., & Allerhand, M. H. (2018). Why do traits come together? The underlying trait and network approaches. In *The SAGE Handbook of Personality and Individual Differences: Volume I: The Science of Personality and Individual Differences*. SAGE Publications Ltd.

Murhadi, W. R., Sutejo, B. S., & X u a ^n, P. T. H. (2024). The big five personality traits Indonesia investor during the covid-19 pandemic. *Media Ekonomi Dan Manajemen*, *39*(1), 99. https://doi.org/10.56444/mem.v39i1.4381

Nasyroh, M., & Wikansari, R. (2017). Hubungan antara kepribadian (big five personality model) dengan kinerja karyawan. *Jurnal Ecopsy*, *4*(1), 10. https://doi.org/10.20527/ecopsy.v4i1.3410

Nugroho, V. T., & Huwae, A. (2025). Neuroticism trait personality, resilience, and loneliness in adolescent victims of bullying. *Journal An-Nafs: Kajian Penelitian Psikologi*, *10*(1), 13–26. https://doi.org/10.33367/psi.v10i1.7023

Pilarska, A. (2018). Big-five personality and aspects of the self-concept: Variable and person-centered approaches. *Personality and Individual Differences*, *127*(January), 107–113. https://doi.org/10.1016/j.paid.2018.01.049

Ritter, A., Esteves, H., & Ferreira, A. (2024). Cultural challenges of the Portuguese adaptation of an IPIP English-based personality inventory. *Psychological Test Adaptation and Development*, *5*(1), 151–162. https://doi.org/10.1027/2698-1866/a000073

Sangwan, N. (2023). Exploring the big five theory: Unveiling the dynamics and dimensions of personality. *Sports Science & Health Advances*, *1*(2), 73–77. https://doi.org/10.60081/SSHA.1.2.2023.73-77

Simanullang, T. (2021). Pengaruh tipe kepribadian the big five model personality terhadap kinerja aparatur sipil negara (kajian studi literatur manajemen keuangan). *Jurnal Manajemen Pendidikan dan Ilmu Sosial*, *2*(2), 747–753. https://doi.org/10.38035/jmpis.v2i2.634

Tarigan, M. (2024). The validity and reliability of the Indonesian IPIP-NEO-120. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, *31*(1).

Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: Early developments, current practice, and future challenges. *Frontiers in Rehabilitation Sciences*, *4*, 1–17. https://doi.org/10.3389/fresc.2023.1208670

Tunç, E. B. (2023). A review of measurement tools developed and adapted based on the rasch model. *Journal of Human and Social Sciences*, *6*(2), 249–275. https://doi.org/10.53048/johass.1369336

Vedel, A., Gøtzsche-Astrup, O., & Holm, P. (2019). The Danish IPIP-NEO-120: A free, validated five-factor measure of personality. *Nordic Psychology*, *71*(1), 62–77. https://doi.org/10.1080/19012276.2018.1470553

Verpalen, A., Van De Vijver, F., & Backus, A. (2018). Bias in dyslexia screening in a Dutch multicultural population. *Annals of Dyslexia*, *68*(1), 43–68. https://doi.org/10.1007/s11881-018-0155-0

Walpuski, M., & Celik, K. N. (2024). Empirical validation of a learning progression in initial chemistry education. *Creative Education*, *15*(2).

Yulisa, K., Widyorini, E., & Primastuti, E. (2023). Hubungan big five personality dan school adjustment dengan agresivitas siswa terhadap guru. *Jurnal Empati*, *12*(6), 424–432. https://doi.org/10.14710/empati.2023.31551