Examining Internal Consistency of the NEO-FFI in an Indian Context: Evidence from Female Undergraduates in Kolkata

Mahammad Abdul Karim

Assistant Professor, Department of Psychology, Sarojini Naidu College for Women, Kolkata- 700 028

KEYWORDS	ABSTRACT
reliability, *Cronbach's alpha,* *NEO-FFI,* *personality assessment*	*Reliability is a cornerstone of psychological testing, ensuring that instruments consistently measure the constructs they intend to capture. The present study examined the internal consistency reliability of the NEO Five-Factor Inventory using Cronbach's alpha among a purposive sample of 101 female undergraduate students in Kolkata, India. The NEO-FFI, a 60-item personality assessment designed to measure the Big Five domains—Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness—was administered, and Cronbach's alpha coefficients were calculated separately for each domain. Results indicated alpha values of .71 for Neuroticism, .60 for Extraversion, .19 for Openness, .61 for Agreeableness, and .71 for Conscientiousness. Compared with the US normative sample of Costa & McCrae, the Indian sample showed consistently lower internal consistency, particularly for Openness. These findings are consistent with prior Indian studies reporting reduced reliability for certain Big Five domains, possibly reflecting cultural orientations, gender-specific response patterns, and limited variance in homogeneous groups. The study underscores the importance of re-examining Western-developed psychometric tools in local contexts and highlights the need for culturally sensitive adaptations of personality inventories.*

1. INTRODUCTION

Psychological research often requires the measurement of abstract constructs such as intelligence, attitudes, emotions, and personality. Unlike physical measurements (e.g., height or weight), these constructs cannot be directly observed; instead, they are inferred through standardized tests and scales. The effectiveness of these instruments depends on two key psychometric

properties: reliability and validity. Reliability refers to the consistency of a measurement, while validity concerns whether the test actually measures what it claims to measure [1]. A reliable instrument ensures that repeated measurements under similar conditions will yield comparable results, thereby reducing the influence of random error. In psychological testing, reliability is not merely a technical requirement but a fundamental condition for meaningful interpretation of scores [2].

One of the most influential frameworks in personality psychology is the Five-Factor Model (FFM), which proposes that human personality can be understood in terms of five broad domains: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness [3]. The NEO Five-Factor Inventory (NEO-FFI), developed by Costa and McCrae[4], is among the most widely used tools to measure these traits. The NEO-FFI is a shortened, 60-item version of the longer NEO Personality Inventory-Revised (NEO-PI-R). Its brevity and accessibility make it popular in both research and applied contexts, including organizational psychology, education, and clinical assessment. Costa and McCrae's [4] original studies reported strong reliability across the five domains, with Cronbach's alpha coefficients ranging from .68 (Agreeableness) to .86 (Neuroticism).

While the NEO-FFI has been extensively validated in Western populations, its application in non-Western and collectivist societies such as India raises important questions about its reliability and cultural sensitivity. Cultural differences influence not only the expression of personality traits but also the way individuals respond to test items [5]. For example, Openness items that emphasize novelty, creativity, and independence may not resonate strongly in collectivist cultures where conformity and tradition are emphasized [6]. Similarly, Agreeableness might manifest differently in hierarchical societies, where deference and compliance are often valued more than egalitarian cooperation. According to Naik and Desai, these differences underscore the importance of examining the psychometric performance of personality scales in diverse cultural contexts [7].

In the Indian context, there is a growing body of literature on the Big Five model, though studies remain fewer compared to Western research. Several Indian studies have reported lower Cronbach's alpha values for Extraversion and Openness compared to the original

Western norms [6][8]. Researchers attribute these findings to cultural values, response styles, and limited variability in homogeneous samples. Furthermore, gender may also play a role in influencing reliability estimates. Studies conducted on female-only samples, for instance, have often shown reduced score variance, which in turn lowers internal consistency estimates (Mehta & Goswami, 2019)[8]. Given that most standardization samples of the NEO-FFI are based on mixed-gender, large, and culturally diverse populations, the applicability of the same reliability expectations to smaller, homogeneous samples in India becomes questionable.

Reliability estimation through Cronbach's alpha is one of the most common practices in psychological research. Cronbach's alpha measures the extent to which items within a test are interrelated and collectively measure the same construct [2]. High alpha values (typically above .70) are generally taken as indicators of good internal consistency [9]. However, alpha is sensitive to several factors, including the number of items, the variability of responses, and cultural relevance of the items. For instance, a domain that resonates poorly with a given cultural group may show low inter-item correlations, thereby producing low alpha even if the construct is theoretically valid. Thus, calculating Cronbach's alpha in different cultural contexts is not merely a statistical exercise but an essential step in establishing the appropriateness of psychological tools.

This study seeks to address these issues by examining the reliability of the NEO-FFI among female college students in Kolkata, India. Specifically, it computes Cronbach's alpha for each of the five personality domains using responses from 101 participants. The choice of this sample is deliberate, as young Indian women in higher education represent a demographic undergoing rapid social and cultural transition—balancing traditional expectations with modern aspirations [10]. Understanding the reliability of personality measures in this group provides valuable insights not only for psychometric validation but also for applied fields such as career counseling, educational interventions, and mental health services.

The study contributes to the growing literature on cross-cultural psychology by:

1. Providing empirical reliability estimates of the NEO-FFI in Indian female student sample.

2. Comparing these results with Costa and McCrae's [4] US norms and other Indian studies.
3. Highlighting cultural and gender-specific factors that may affect psychometric performance.

By doing so, this research underscores the importance of cultural adaptation in psychological assessment and encourages cautious interpretation of Western-developed scales in non-Western contexts.

2. LITERATURE REVIEW

2.1 Reliability in Psychological Testing

Reliability is one of the fundamental pillars of psychological measurement. According to Classical Test Theory (CTT), an observed score (X) is a combination of a true score (T) and an error component (E), expressed as $X = T + E$ [11]. Reliability refers to the degree to which the observed score consistently reflects the true score across repeated measurements. If error variance is high, reliability is low; conversely, when error variance is minimized, reliability is maximized.

Among the various methods of assessing reliability, internal consistency is particularly crucial when dealing with personality inventories, which rely on multiple items to represent broad latent traits. Cronbach's alpha [2] is the most widely used measure of internal consistency. It reflects the average inter-item correlation adjusted for the number of items in a scale. While values above 0.70 are considered acceptable for research purposes [9], the interpretation of alpha must be contextualized—shorter scales, homogeneous samples, and culturally incongruent items often yield lower alpha values.

2.2 The Five-Factor Model of Personality and the NEO-FFI

The Five-Factor Model (FFM) of personality has emerged as a dominant paradigm in personality psychology [3]. It posits that personality can be summarized into five broad traits: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Each domain represents a continuum of behavioral and emotional tendencies, from emotional instability to stability (Neuroticism), sociability to introversion (Extraversion), curiosity to conventionality (Openness), cooperativeness to antagonism (Agreeableness), and self-discipline to impulsiveness (Conscientiousness).

To operationalize this model, Costa and McCrae developed the NEO Personality Inventory-Revised (NEO-PI-R) and its shorter version, the NEO Five-Factor Inventory (NEO-FFI). The NEO-FFI consists of 60 items, with 12 items per domain, rated on a 5-point Likert scale. While the NEO-PI-R offers greater depth, the NEO-FFI is valued for its brevity, making it practical for use in large-scale research and applied contexts such as education and organizational psychology.

2.3 Reported Reliability of the NEO-FFI in Western Contexts

Costa and McCrae [4] reported high internal consistency coefficients for the NEO-FFI based on a US normative sample of 1000 participants (500 males, 500 females), aged 21–96 years. Their findings indicated Cronbach's alphas of .86 for Neuroticism, .77 for Extraversion, .73 for Openness, .68 for Agreeableness, and .81 for Conscientiousness. These values established the NEO-FFI as a psychometrically robust tool for measuring the Big Five traits.

Subsequent studies across diverse Western populations have largely replicated these reliability findings. For instance, Robins et al. [12] reported alphas ranging from .70 to .85 in a large college sample. McCrae and Terracciano [13], analyzing observer ratings across 50 cultures, confirmed the generalizability of the FFM, although they acknowledged some variability in Openness and Agreeableness.

Nevertheless, even in Western samples, Openness often demonstrates lower reliability than the other domains, reflecting its conceptual breadth and cultural sensitivity [14]. This observation foreshadows the greater challenges encountered in non-Western adaptations.

2.4 Cross-Cultural Adaptations of the NEO-FFI

The globalization of psychology has prompted extensive cross-cultural research on the FFM. Studies across Europe, Asia, and Africa have generally supported the broad structure of the model but highlighted cultural nuances in reliability and validity [5][13].

In collectivist cultures, Extraversion and Openness often yield lower alpha values. For example, Yamagata et al. [15] in Japan and Zhang et al. [16] in China found that participants tended to respond less variably to Openness items, reflecting cultural emphasis on conformity and tradition. Similarly, collectivist values can influence Agreeableness, where compliance and deference may be more strongly emphasized than cooperation in egalitarian terms [17].

Cross-cultural research thus highlights the importance of cultural adaptation rather than mere translation. Items that capture the essence of a trait in one culture may fail to evoke meaningful variation in another, leading to lower inter-item correlations and reduced reliability.

2.5 Reliability Studies of the NEO-FFI in India

The Indian context offers a rich yet underexplored terrain for personality research. With its cultural emphasis on collectivism, hierarchical social structures, and rapid modernization, India presents a unique backdrop for evaluating the NEO-FFI.

Naik and Desai [7] examined the Big Five traits among Indian undergraduates and reported that Neuroticism and Conscientiousness demonstrated acceptable reliability, while Openness and Agreeableness showed weaker coefficients compared to Western norms. They attributed this to cultural influences, where traits like intellectual curiosity and novelty-seeking (central to Openness) may be undervalued in traditional educational contexts.

Singh and Choudhary [6] conducted a psychometric evaluation of the NEO-FFI in Indian student populations and observed notably lower alphas for Extraversion and Openness. They argued that cultural tendencies toward modesty and social restraint could reduce variability in responses, thereby lowering internal consistency.

Mehta and Goswami [8] highlighted the role of gender-specific samples, noting that female groups often display reduced variance in certain traits due to shared socialization experiences, which may deflate reliability coefficients. This observation is particularly relevant to the present study, which focuses exclusively on female college students.

Other Indian studies echo similar trends. Sharma and Singh [18] found moderate reliability for Neuroticism and Conscientiousness but weaker performance for Openness and Agreeableness among postgraduate students. Rao and Suneetha [19] emphasized the need for localized adaptations of Western-developed instruments, arguing that literal translations may not adequately capture the socio-cultural nuances of personality expression in India.

Taken together, these findings indicate a consistent pattern: while Neuroticism and Conscientiousness are relatively stable across cultural contexts, Openness and Extraversion appear more culturally contingent in India.

2.6 The Case of Openness in India

Among the five domains, Openness to Experience is the most problematic in terms of reliability in Indian samples. Openness encompasses facets such as imagination, curiosity, aesthetic appreciation, and willingness to explore novelty. However, in collectivist contexts like India, conformity to social norms, respect for tradition, and emphasis on group harmony often limit the overt expression of these traits [20].

For example, a student who scores low on items measuring creativity or novelty-seeking may not necessarily lack those qualities but may have been socialized to suppress such tendencies. This cultural misalignment reduces inter-item correlations, resulting in low Cronbach's alpha values. Indeed, the extremely low alpha (.19) for Openness in the present study mirrors earlier findings [6][7].

2.7 Importance of Gender-Homogeneous Samples

Gender composition also affects reliability estimates. Homogeneous samples, particularly all-female groups, often display less variability in certain domains [8]. For instance, in patriarchal societies, young women may share similar experiences of socialization that emphasize modesty, compliance, and relational harmony. Such uniformity reduces item variance, which is critical for reliability. Thus, while gender-homogeneous samples are useful for focused analysis, they may inherently yield lower alpha values compared to mixed-gender or more diverse groups.

2.8 Summary of Literature Gaps

The literature reveals three key gaps that the present study seeks to address:

1. Limited Indian research on the NEO-FFI – While a handful of studies have examined its psychometric properties in India, most focus on small or region-specific samples. Broader evidence is needed to establish generalizable reliability benchmarks.

2. Underrepresentation of female-only samples – Most studies combine genders, leaving a gap in understanding how personality measures perform in gender-homogeneous groups.

3. Insufficient focus on cultural adaptation – Many Indian studies use literal translations without addressing cultural equivalence, which may contribute to poor reliability in domains like Openness.

By analyzing Cronbach's alpha values of the NEO-FFI among female college students in Kolkata, this study contributes to filling these gaps. It not only provides updated reliability estimates but also contextualizes them within cultural and gender-specific frameworks.

3. METHODOLOGY

3.1 Research Design

The present study employed a quantitative, descriptive research design aimed at evaluating the internal consistency reliability of the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae [4] using Cronbach's alpha. As the central focus was to determine the psychometric soundness of the NEO-FFI in an Indian context, particularly among female college students in Kolkata, the design was restricted to the single administration of the test, followed by computation of item-level and domain-level reliability estimates.

Reliability studies of this nature are non-experimental, as they do not involve manipulation of independent variables but rather focus on the measurement properties of an instrument [1]. This design allows for the systematic collection of quantitative data, ensuring replicability and transparency in statistical procedures.

3.2 Participants

The sample comprised 101 female undergraduate students recruited from Kolkata, West Bengal.

Inclusion criteria were:

1. Female students enrolled in undergraduate courses,
2. Age range between 18 and 22 years,
3. Proficiency in English sufficient to understand the NEO-FFI items, and
4. Voluntary consent to participate in the study.

Exclusion criteria included students with known psychological or neurological disorders that could interfere with test-taking, as well as incomplete or invalid responses.

The mean age of participants was 19.8 years (SD = 1.21). Socioeconomic backgrounds varied, though the majority were from urban, middle-class families. The choice of an all-female group was intentional to explore gender-homogeneous reliability estimates, given prior evidence that gender composition influences psychometric outcomes [8].

Sampling was carried out using a purposive sampling strategy, which is common in reliability and validation studies where the focus is on psychometric evaluation rather than general population inference [6].

3.3 Instrument

The NEO Five-Factor Inventory of Costa & McCrae [4] was used. It is a 60-item personality inventory that assesses the five major domains of personality:

1. Neuroticism (N) – 12 items measuring emotional stability versus instability.
2. Extraversion (E) – 12 items measuring sociability, activity, and positive affect.
3. Openness to Experience (O) – 12 items measuring imagination, curiosity, and aesthetic appreciation.
4. Agreeableness (A) – 12 items measuring interpersonal trust, altruism, and cooperation.
5. Conscientiousness (C) – 12 items measuring orderliness, responsibility, and self-discipline.

Responses were recorded on a 5-point Likert scale ranging from 0 (Strongly Disagree) to 4 (Strongly Agree). Each domain score is derived by summing responses across the 12 relevant items.

The NEO-FFI was originally developed and validated in the United States and has been used extensively in cross-cultural research. Previous Indian studies [7] have reported mixed reliability outcomes, particularly for Openness and Extraversion, making it suitable for reassessment in a fresh sample.

3.4 Materials Required

1. Test booklets of the NEO-FFI,
2. Demographic questionnaire (age, course, year of study, socioeconomic status),
3. Informed consent form

3.5 Procedure

The study was conducted in three major phases: preparation, administration, and computation.

Phase 1: Preparation

Permission was sought from college authorities to conduct the study during scheduled hours. Participants were briefed about the study's purpose, assured of confidentiality, and informed that participation was voluntary. Written informed consent was collected.

Phase 2: Administration

The NEO-FFI was administered in a group setting, with approximately 20–25 students per session in classroom-like conditions. Standardized instructions were read aloud to minimize experimenter bias. Participants were asked to respond honestly and were

assured that there were no right or wrong answers. Average completion time was 12–15 minutes. All participants completed the test without interruptions.

Phase 3: Data Entry and Computation

Completed questionnaires were screened for missing responses. Data were entered into Microsoft Excel. Item-level variances ($\Sigma\sigma_i^2$) were computed separately for each domain, as well as the total variance of summed domain scores (σ^2). Cronbach's alpha was then calculated using the formula (Cronbach, 1951):

$$\propto = \frac{n}{n-1}\left(1 - \frac{\Sigma(\sigma_i^2)}{\sigma^2}\right)$$

Where:
- n = number of items in a domain (12 for each Big Five trait),
- $\Sigma\sigma_i^2$ = sum of item variances, and
- σ^2 = variance of total test scores for that domain.

3.6 Data Analysis

The main analysis focused on computing Cronbach's alpha for each of the five domains. These values were then compared with reference values reported by Costa and McCrae [4] in their US normative sample.

Interpretation followed commonly accepted benchmarks of George & Mallery [21] :
- ≥ 0.90 = Excellent reliability,
- 0.80–0.89 = Good,
- 0.70–0.79 = Acceptable,
- 0.60–0.69 = Questionable,
- 0.50–0.59 = Poor,
- < 0.50 = Unacceptable.

Additionally, differences between Indian and US samples were analyzed qualitatively in terms of cultural factors, gender composition, and sample size.

3.7 Rationale for Methodological Choices

The decision to focus on an all-female group was deliberate, addressing a gap in Indian reliability studies that typically include mixed samples. Using Excel for calculations ensured transparency and replicability, while group administration reduced time constraints. However, the choice of purposive sampling and the limited regional scope (Kolkata) represent methodological trade-offs, restricting generalizability but enabling focused reliability evaluation.

4. RESULTS

The primary aim of this study was to evaluate the internal consistency reliability of the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae) [4] in an Indian context, specifically among female undergraduate students in Kolkata. Reliability was assessed using Cronbach's alpha coefficients for each of the five domains: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Results are presented in two stages: (a) alpha values for the present sample, and (b) comparison with the normative values reported in the US population.

4.1 Cronbach's Alpha for the Indian Sample

Table 1 shows the Cronbach's alpha coefficients calculated from the responses of 101 female students.

Table 1. Cronbach's Alpha for Indian Sample (N = 101, Female, Kolkata)

Domain	Cronbach's Alpha
Neuroticism	0.71
Extraversion	0.60
Openness	0.19
Agreeableness	0.61
Conscientiousness	0.71

Interpretation of these results using George and Mallery's [21] guidelines suggests that:
- Neuroticism (α = 0.71) and Conscientiousness (α = 0.71) fall within the *acceptable* range.
- Extraversion (α = 0.60) and Agreeableness (α = 0.61) lie in the *questionable* range.
- Openness (α = 0.19) is unacceptably low, indicating severe problems with internal consistency in this sample.

4.2 Comparison with US Normative Sample

For a meaningful interpretation, the coefficients were compared with the original reliability estimates reported by Costa and McCrae [4] in their US general population sample (N = 1000; 500 males, 500 females; age range 21–96 years).

Table 2. Cronbach's Alpha: Indian Sample vs. US Normative Sample

Domain	Indian Sample (N = 101, Female)	US Normative Sample (N = 1000)
Neuroticism	0.71	0.86
Extraversion	0.60	0.77
Openness	0.19	0.73
Agreeableness	0.61	0.68

Conscientiousness	0.71	0.81

The comparison reveals that all domains in the Indian sample show lower reliability coefficients than those reported in the US normative data. The most striking difference lies in the domain of Openness, where the Indian sample's alpha of 0.19 is drastically lower than the US value of 0.73.
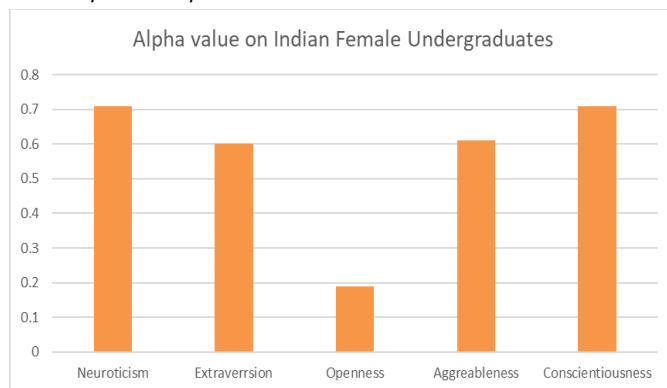
4.3 Graphical Representation

Alpha value on Indian Female Undergraduates

Figure 1

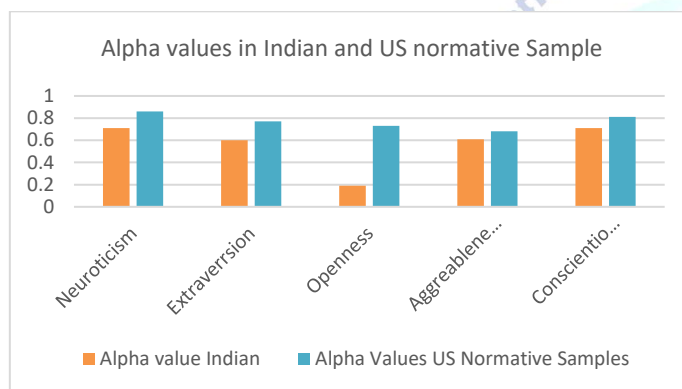Alpha values in Indian and US normative Sample

Figure 2

4.4 Cross-Cultural Insights

The results suggest that Western-developed personality measures may not transfer seamlessly to Indian contexts. Several possible explanations we can infer:

- Cultural variation in trait salience: Traits like Openness may not be strongly emphasized in collectivist, exam-oriented cultures .
- Gender homogeneity: The all-female sample may have produced reduced variance in responses, particularly in traits influenced by social norms (Mehta & Goswami, 2019)[8].
- Sample size effects: With only 101 participants, reliability coefficients are less stable than those computed from the large US normative sample of 1000 (Cortina) [9].

- Language and comprehension issues: Although the NEO-FFI was administered in English, subtle nuances in item wording may have reduced clarity or cultural relevance for Indian college students (Sinha & Tripathi)[22].
- The domains of Neuroticism and Conscientiousness achieved acceptable internal consistency.
- Extraversion and Agreeableness demonstrated questionable but interpretable reliability.
- Openness displayed extremely poor reliability, raising concerns about its cross-cultural applicability in the present context.

These results highlight the necessity of revising, adapting, or culturally validating certain items within the NEO-FFI before its use in Indian educational or psychological settings.

5. DISCUSSION

The present study investigated the internal consistency reliability of the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992)[4] among a group of 101 female undergraduate students from Kolkata, India. Reliability was assessed using Cronbach's alpha for each of the five domains: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. The results revealed variability across domains, with Neuroticism (α = 0.71) and Conscientiousness (α = 0.71) showing acceptable reliability, Extraversion (α = 0.60) and Agreeableness (α = 0.61) demonstrating questionable reliability, and Openness (α = 0.19) reflecting unacceptably low reliability.

5.1 Comparison with Normative and International Findings

When compared to the US normative sample reported by Costa and McCrae (1992)[4], it is evident that the Indian sample consistently displayed lower reliability coefficients. In the US context, alphas ranged from 0.68 (Agreeableness) to 0.86 (Neuroticism), whereas the present study's values ranged from 0.19 to 0.71. The starkest contrast was observed in the Openness domain, where reliability dropped from 0.73 (US sample) to 0.19 (Indian female sample).

Similar cross-cultural differences have been reported in prior research. McCrae et al. [13] found that Openness and Agreeableness often varied in reliability across non-Western samples, particularly in collectivist societies. Studies conducted in India [6][8] also reported weaker internal consistency in Openness and

Extraversion, suggesting that the measurement of certain Big Five traits may not fully align with Indian cultural and educational contexts.

5.2 Explaining Domain-Specific Differences

5.2.1 Neuroticism

The relatively acceptable reliability for Neuroticism (α = 0.71) aligns with the notion that emotional stability and vulnerability are universally relevant constructs (Costa & McCrae) [4]. Indian students may experience consistent patterns of stress, anxiety, and mood fluctuations due to academic and social pressures (Arora & Singh)[23], which could explain the moderate internal consistency.

5.2.2 Conscientiousness

Conscientiousness (α = 0.71) also showed acceptable reliability. Traits such as organization, responsibility, and discipline are reinforced in academic contexts in India, where performance in competitive examinations and structured learning is emphasized [24]. The stable cultural value placed on industriousness likely contributes to more consistent responses.

5.2.3 Extraversion

Extraversion (α = 0.60) was found to be questionable. Indian collectivist norms often discourage overt displays of sociability, assertiveness, and spontaneity among young women [25]. Thus, the inconsistent responses might reflect tension between cultural expectations and the Western conceptualization of Extraversion.

5.2.4 Agreeableness

The Agreeableness domain (α = 0.61) yielded a questionable reliability level. While interpersonal harmony and cooperation are emphasized in collectivist societies (Chaudhary & Sharma)[26], differences in interpretation of items (e.g., "forgiving others easily" vs. "standing up for oneself") could lead to varied responses. The alpha may be deflated due to the multidimensional nature of Agreeableness as it manifests in Indian culture.

5.2.5 Openness

The Openness domain (α = 0.19) displayed critically low internal consistency. This aligns with prior Indian studies where Openness consistently showed weaker psychometric properties (Singh & Choudhary; Naik & Desai)[6][7]. Several explanations are possible:

- Cultural orientation: Indian educational settings often emphasize conformity, rote learning, and adherence to traditions rather than intellectual curiosity or novelty-seeking (Sinha & Tripathi) [22].
- Homogeneity of the sample: Since the sample consisted only of female college students, variance in responses was limited, reducing alpha.

Item relevance: Some items measuring openness to aesthetics, liberal values, or abstract thinking may not resonate strongly with participants, leading to inconsistent responses

5.3 Gender and Cultural Contexts

The all-female composition of the sample is an important factor. Research suggests that women often respond more homogeneously on personality measures, especially in domains like Openness and Agreeableness (Mehta & Goswami)[8]. Furthermore, gender norms in India may restrict the expression of certain traits, such as assertiveness (Extraversion) or non-conformity (Openness), which contributes to weaker reliability.

Cultural factors also play a major role. The Big Five model, though considered universal, has been critiqued for its Western-centric development (Cheung et al)[5]. In collectivist contexts such as India, alternative models—such as the Indian Personality Inventory [27]—sometimes yield more culturally valid results.

6. CONCLUSION

6.1 Conclusion

This study assessed the internal consistency reliability of the NEO-FFI among 101 female college students in Kolkata, India. Cronbach's alpha values ranged from very low (Openness = .19) to moderate (Neuroticism and Conscientiousness = .71). Compared to the original US normative data (Costa & McCrae)[4], reliability scores were generally lower, especially for Openness.

The findings reinforce the importance of validating Western psychological instruments in diverse cultural contexts before applying them in research or clinical practice. As Indian psychology increasingly engages with global psychometrics, it becomes essential to balance imported tools with locally adapted or indigenous measures (Srivastava; Tripathi & Kapoor) [28][29].

The study concludes that while the NEO-FFI demonstrates acceptable reliability for certain traits in the Indian context, significant caution is warranted, particularly for Openness. Without cultural adaptation,

reliance on raw NEO-FFI scores may lead to misleading inferences about personality in Indian populations.

6.2 Implications of the Findings

The results highlight critical insights for psychological testing, education, and applied research in India:

1. Cross-cultural differences: Traits like Openness and Extraversion may be conceptualized and expressed differently in collectivist societies compared to Western contexts (McCrae et al.)[13]. For example, creativity and nonconformity (central to Openness) may not be equally encouraged within traditional or exam-oriented educational systems in India.

2. Gendered expectations: The use of an all-female sample highlights how cultural gender norms might affect reliability. Indian women may respond more uniformly to items related to conformity, interpersonal harmony, and cautiousness, which may reduce response variability and lower internal consistency (Mehta & Goswami)[8].

3. Psychological practice: In applied contexts, such as counseling, career guidance, and organizational recruitment, practitioners should be cautious in using the NEO-FFI without adaptation. Over-reliance on raw scores from domains with low reliability (e.g., Openness) may lead to misinterpretation of traits.

4. Educational utility: Since personality assessments are increasingly used in higher education for career planning and personal development, there is a need to ensure that the instruments employed are psychometrically sound for local populations [7].

6.3 Limitations of the Study

While this study contributes to the body of literature, several limitations must be acknowledged:

- Restricted sample composition: Only female college students were included. This limits the generalizability of findings across genders, age groups, and educational backgrounds.

- Sample size: The study utilized 101 participants, which, although adequate for preliminary reliability estimation, is small compared to normative studies of the NEO-FFI (Costa & McCrae) [4].

- Geographic limitation: The participants were drawn exclusively from Kolkata, which may not represent other cultural regions of India, such as rural areas or different linguistic states.

- Single reliability estimate: Only Cronbach's alpha was calculated. Other metrics such as McDonald's Omega, test-retest reliability, or split-half reliability were not explored.

- Self-report limitations: The NEO-FFI, being a self-report inventory, is vulnerable to social desirability bias, particularly in collectivist cultures where modesty and conformity are valued (Sinha & Tripathi) [22].

6.4 Scope for Future Research

The study opens several avenues for future investigations:

1. Larger and more diverse samples: Future research should include larger samples across gender, socioeconomic status, and regional backgrounds in India. Inclusion of rural, semi-urban, and urban populations will allow for more representative findings.

2. Cross-gender comparisons: Examining male, female, and mixed-gender groups will help determine whether the observed reliability differences are specific to female participants or extend more broadly across populations.

3. Cultural adaptation of items: Specific items in Openness and Extraversion may require translation, rewording, or cultural modification to better capture relevant expressions of these traits in India. For instance, instead of focusing solely on liberal values or aesthetic appreciation, items could be reframed around adaptability, practical creativity, or community-based innovations.

Comparative studies with indigenous tools: Comparative analyses of the NEO-FFI with Indian-developed instruments, such as Sinha's Indian Personality Inventory [27], could help establish convergent validity and highlight cultural differences in trait measurement.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Anastasi, A., & Urbina, S. (1997). Psychological testing (7th ed.). Prentice Hall.

[2] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334. https://doi.org/10.1007/BF02310555

[3] McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. Journal of Personality, 60(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

[4] Costa, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Psychological Assessment Resources.

[5] Cheung, F. M., van de Vijver, F., & Leong, F. T. L. (2011). Toward a new approach to the study of personality in culture. American Psychologist, 66(7), 593–603. https://doi.org/10.1037/a0022389

[6] Singh, A., & Choudhary, P. (2015). Psychometric evaluation of NEO-FFI in Indian student populations. Journal of Indian Academy of Applied Psychology, 41(3), 234–242.

[7] Naik, N., & Desai, G. (2017). Big Five personality traits and academic achievement among Indian college students. Indian Journal of Positive Psychology, 8(2), 134–138.

[8] Mehta, S., & Goswami, T. (2019). Reliability and cultural adaptation of Big Five personality measures in India. Indian Journal of Psychology, 94(2), 45–57.

[9] Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 78(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

[10] Nisbett, R. E. (2003). The geography of thought: How Asians and Westerners think differently...and why. Free Press.

[11] Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). McGraw-Hill.

[12] Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality correlates of self-esteem. Journal of Research in Personality, 35(4), 463–482.

[13] McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. Journal of Personality and Social Psychology, 88(3), 547–561. https://doi.org/10.1037/0022-3514.88.3.547

[14] McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae & J. Allik (Eds.), The five-factor model of personality across cultures (pp. 105–125). Kluwer Academic/Plenum.

[15] Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., … Jang, K. L. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study. Journal of Personality and Social Psychology, 90(6), 987–998. https://doi.org/10.1037/0022-3514.90.6.987

[16] Zhang, L., Mandl, H., & Wang, E. (2011). Personality, self-concept and academic achievement: An investigation of Chinese university students. Educational Psychology, 31(6), 721–729. https://doi.org/10.1080/01443410.2011.587709

[17] Heine, S. J. (2016). Cultural psychology (3rd ed.). W.W. Norton

[18] Sharma, R., & Singh, S. (2018). Reliability and validity of NEO-FFI among postgraduate students in India. Indian Journal of Research in Psychology, 12(2), 44–52.

[19] Rao, N., & Suneetha, A. (2019). Psychometric evaluation of Western-developed personality inventories in Indian samples: A systematic review. Indian Journal of Psychological Science, 10(1), 88–99

[20] Roland, A. (1988). In search of self in India and Japan: Toward a cross-cultural psychology. Princeton University Press.

[21] George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference (4th ed.). Allyn & Bacon.

[22] Sinha, J. B. P., & Tripathi, R. C. (1994). Individualism in a collectivist culture: A case of coexistence of opposites. International Journal of Psychology, 29(6), 729–738

[23] Arora, R., & Singh, S. (2019). Academic stress and personality traits among Indian college students. Indian Journal of Behavioral Sciences, 25(1), 12–23.

[24] Basu, S. (2017). Education, competition and student stress in India: A sociocultural analysis. Journal of Indian Psychology, 35(2), 55–67.

[25] Gupta, R., & Sharma, M. (2016). Extraversion, social norms, and cultural expectations: A study among Indian college women. Journal of Personality and Social Behavior, 14(3), 199–210

[26] Chaudhary, A., & Sharma, R. (2017). Cultural influences on agreeableness: A study of collectivist values in Indian students. Indian Journal of Psychology and Education, 7(2), 45–56.

[27] Sinha, J. B. P. (2002). Indian Personality Inventory (IPI): Manual. Agra Psychological Research Cell.

[28] Srivastava, S. (2010). Measuring the Big Five personality factors. Indian Journal of Psychology and Education, 1(1), 23–32.

[29] Tripathi, R., & Kapoor, A. (2017). Cross-cultural validation of Big Five personality inventories in India: A systematic review. Indian Journal of Psychological Science, 8(2), 221–230.