

Development and Preliminary Validation of Psychometric Assessment Tools for Educational Stream Selection among Class 10th Students

**Dr. Prabhdeep Singh Chawla¹, Ms. Gunpreet Kaur²,
Ms. Prabhsimran Kaur³**

¹Project Coordinator, Council for Vocational Training and Education, Punjab

²Proprietor, Education, Math Vibes

³Student, Education, Chandigarh University

Abstract

This study addressed a critical gap in Indian educational psychology, asked a simple but important question can we measure personality, aptitude and career interest accurately in Class 10th students studying in schools ? the absence of locally validated personality assessment tools for adolescent students. A 50-item Big Five Personality Scale was developed and validated specifically for Class 10 students (ages 15–16) in government schools across Punjab, India, measuring Conscientiousness, Openness to Experience, Neuroticism, Agreeableness, and Extraversion (the OCEAN framework).

The study followed a two-phase design. In Phase 1, the complete 50-item scale was administered to N = 521 Class 10th students (after list wise deletion) from 17 schools across urban, semi-urban, and rural settings. Exploratory Factor Analysis (EFA) using Principal Axis Factoring with Varimax rotation identified 10 items with systematic factor migration; these were removed, yielding a 40-item scale (8 items per trait). In Phase 2, the revised scale was administered to an independent sample of N = 101 students for reliability cross-validation.

Phase 1 EFA (loading criterion $\geq .35$) produced the following item classification across the 40 retained items: 32 items loaded correctly and strongly on their intended factor ($\geq .35$); 1 item (Q47, Openness) loaded on its correct factor but below the loading threshold ($\lambda = .271$); and 7 items showed factor migration to a non-intended factor. Together, 8 items (1 weak-loading + 7 migrated) are flagged for future wording revision; 32 + 1 + 7 = 40 items in total. Primary factor agreement rate (items on correct factor, regardless of strength) was 33/40 (82.5%). The five-factor solution explained 41.6% of total variance.

Phase 2 internal consistency improved across all five traits to the Good range (Cronbach's $\alpha = .827-.894$; full-scale $\alpha = .879$). Item-total correlations ranged from .478 to .833. All 40 items showed Good to Excellent discrimination indices (all $p < .001$). KMO = .776 (Middling) and Bartlett's Test $p < .001$ confirmed factor analytic suitability.

The 40-item scale is established as a psychometrically sound instrument for personality assessment of Class 10 students in Punjab government schools, with implications for school counselling and career guidance discussed.

Keywords: Big Five personality, psychometric validation, Class 10 students, Punjab, Cronbach's alpha, Principal Axis Factoring, Varimax rotation, item discrimination, school counselling, adolescent assessment, OCEAN

1. INTRODUCTION

Every student who walks into a classroom brings with them not just a textbook and a pencil, but a whole personality — a set of tendencies that shape how they study, how they get along with others, how they handle stress, and what kind of future they might be moving towards. Understanding personality is not just an academic exercise. For teachers, counsellors, and the students themselves, it can be genuinely useful information.

The Big Five model of personality is the most widely tested and respected framework for describing human personality in psychology today. Researchers across dozens of countries and languages have found that five broad dimensions — Conscientiousness, Openness to Experience, Neuroticism, Agreeableness, and Extraversion — capture most of what is meaningfully different from one person to another (McCrae & Costa, 1987; Goldberg, 1990). These five traits have been linked to real outcomes including academic achievement, career satisfaction, mental health, and the quality of relationships.

Almost all well-known Big Five questionnaires were designed and validated in Western countries. When applied in India — especially with adolescents in government schools — there is no guarantee that items will be understood in the same way or that scores will carry equivalent meaning. As Srivastava and Misra (2019) noted, school counsellors in India largely work with borrowed tools whose suitability for Indian students has never been properly established.

Class 10 is one of the most significant moments in an Indian student's life. Board examinations, stream selection, and the first real conversations about the future all happen at this stage. A personality questionnaire properly validated for this group — in this language, in this cultural setting — could support counsellors in having more informed, evidence-based conversations with students. That is what this study set out to build.

1.1 Why This Study Was Needed

School counsellors in Punjab work within a high-stakes educational context shaped by competitive Board examinations, strong family expectations around academic performance, and collectivist social values that influence how adolescents experience and express personality. Instruments adapted from individualistic Western cultures — where assertiveness is framed positively, emotional expression is more individually calibrated, and the relationship between discipline and anxiety is weaker — may not function equivalently in this context (Singh & Bhargava, 2015; Kaur & Verma, 2021).

Research consistently shows that Conscientiousness predicts academic success (Poropat, 2009) and Neuroticism is associated with examination anxiety — both highly relevant for Class 10 students. But these connections can only be meaningfully studied and applied if the measurement tool works correctly for this population. This study is an attempt to make that possible.

1.2 Objectives

The study had the following specific aims:

1. To develop a 50-item Big Five Personality Scale for Class 10 students, with items arranged in an interleaved 1-in-5 rotation to reduce sequential response sets.
2. To assess item quality through corrected item-total correlations and Cronbach's alpha-if-item-deleted analysis.

3. To evaluate item discrimination using the upper-lower 27% method with independent samples t-tests.
4. To assess factor analytic suitability through the KMO measure and Bartlett's Test of Sphericity.
5. To conduct EFA (Principal Axis Factoring, Varimax rotation, five factors a priori) and evaluate structural validity using a loading criterion of $\geq .35$.
6. To identify and remove systematically migrating items, yielding a refined 40-item scale.
7. To cross-validate internal consistency in a second independent sample ($N = 101$).
8. To examine inter-trait correlations as evidence of discriminant and convergent validity.

2. REVIEW OF LITERATURE

2.1 The Big Five Model

The Big Five framework emerged from lexical research into personality-descriptive language (Allport & Odbert, 1936; Cattell, 1943; Tupes & Christal, 1961) and was formalized by Costa and McCrae (1985, 1992) and Goldberg (1990). The five dimensions — Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism (OCEAN) — have been replicated across dozens of languages and cultures (McCrae et al., 2005), and each dimension has demonstrated consistent predictive validity for educational and occupational outcomes.

Conscientiousness consistently predicts academic performance across educational levels, often as strongly as intelligence (Poropat, 2009). Neuroticism predicts examination anxiety. Extraversion is associated with classroom participation and peer relationships. Openness predicts intellectual curiosity and creative engagement. These connections make the Big Five genuinely useful in educational settings, not just in research laboratories.

2.2 Personality in Indian Educational Settings

Despite the global evidence base for the Big Five, systematic validation studies with Indian adolescents remain limited. Most personality tools used in Indian schools were developed elsewhere and adapted primarily through translation rather than full psychometric revalidation (Lodhi et al., 2002). Several researchers have noted that item content designed for individualistic Western cultures may not function equivalently in Indian contexts shaped by collectivist values, joint family structures, and different social expectations around emotional expression and assertiveness (Singh & Bhargava, 2015; Kaur & Verma, 2021).

Mathur and Gupta (2020) found that the relationship between Conscientiousness and Neuroticism tends to be closer in Indian adolescent samples than in Western ones — a finding replicated in the present study and discussed in Section 5.4.

2.3 Psychometric Standards for Scale Validation

Cronbach's alpha is the most common index of internal consistency: $\geq .70$ = acceptable, $\geq .80$ = good, $\geq .90$ = excellent (George & Mallery, 2003; Nunnally, 1978). The standard EFA approach uses Principal Axis Factoring (PAF) as the extraction method and Varimax rotation to produce an interpretable orthogonal solution (Fabrigar et al., 1999). Primary factor loadings $\geq .35$ are generally considered meaningful (Comrey & Lee, 1992; Stevens, 2002). The KMO measure ($\geq .60$) and Bartlett's Test ($p < .001$) serve as required prerequisites for EFA (Kaiser, 1974). Discrimination indices $\geq .40$ are considered excellent (Ebel & Frisbie, 1991).

3. METHOD

3.1 Research Design

The study followed a two-phase instrument development and cross-validation design (Clark & Watson, 1995; Worthington & Whittaker, 2006). Phase 1 served as the main development phase; Phase 2 was an independent reliability cross-validation.

3.2 Participants and Sample Characteristics

Phase 1

A total of 525 Class 10 students from 17 government schools across Punjab were recruited. After list wise deletion of incomplete records, N = 521 students provided complete data on the personality subscales (the EFA was conducted on N = 520, those with complete data on all 40 retained items). The sample comprised 295 girls (56.6%) and 226 boys (43.4%). Schools were drawn from urban (n = 147, 28.2%), semi-urban (n = 166, 31.9%), and rural (n = 208, 39.9%) settings. All students were approximately 15–16 years old and native speakers of Punjabi or Hindi. Table 1 presents full sample characteristics. Participation was voluntary and school authorities approved the study.

Phase 2

N = 101 Class 10 girls from government schools across urban (n = 40, 39.6%), semi-urban (n = 30, 29.7%), and rural (n = 31, 30.7%) settings participated in Phase 2. All 101 students provided complete data on the 40-item revised scale. The Phase 2 sample was drawn from girls' sections of the participating schools; all participants were female. Phase 2 was designed for reliability cross-validation only; its N:item ratio (101:40 = 2.5:1) is below the minimum for stable EFA (MacCallum et al., 1999; minimum = 10:1), and no EFA was conducted.

Table 1
Sample Characteristics — Phase 1 (N = 521) and Phase 2 (N = 101)

Characteristic	Category	Phase 1	Phase 2
Total sample	Students	N = 521	N = 101
Gender	Female	295 (56.6%)	101 (100%)
	Male	226 (43.4%)	—
School locality	Urban	147 (28.2%)	40 (39.6%)
	Semi-urban	166 (31.9%)	30 (29.7%)
	Rural	208 (39.9%)	31 (30.7%)
School type	Government	17 schools	Same school setting
Age range	15–16 years	521 (100%)	101 (100%)
Medium of instruction	Punjabi or Hindi	521 (100%)	101 (100%)
Class	Class 10th (Board year)	521 (100%)	101 (100%)

3.3 The Questionnaire

The scale comprised 50 statements, each rated on a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). Each personality trait was measured by 10 items in an interleaved 1-in-5 rotation: Conscientiousness — Items 1, 6, 11, 16, 21, 26, 31, 36, 41, 46; Openness — Items 2, 7, 12, 17, 22, 27, 32, 37, 42, 47; Neuroticism — Items 3, 8, 13, 18, 23, 28, 33, 38, 43, 48; Agreeableness — Items 4, 9, 14, 19, 24, 29, 34, 39, 44, 49; Extraversion — Items 5, 10, 15, 20, 25, 30, 35, 40, 45, 50. Items were written in simple, culturally accessible language and reviewed by two educational psychologists prior to administration.

3.4 Procedure

The questionnaire was administered on paper in classrooms during school hours by trained research assistants. Standardised oral instructions were delivered in Hindi or Punjabi. Participants were told there were no right or wrong answers, responses were confidential, and results had no bearing on academic records. Completion time was approximately 20–25 minutes. For Phase 2, the 10 removed items were marked "Not Applicable" and left blank, retaining original item numbering for traceability.

3.5 Statistical Analyses

The following analyses were performed:

1. Descriptive statistics (M, SD, range) at item and trait levels.
2. Corrected item-total correlations and Cronbach's alpha-if-item-deleted.
3. Cronbach's alpha (internal consistency) and Spearman-Brown corrected split-half reliability (odd-even split).
4. Upper 27% vs Lower 27% discrimination indices with independent samples t-tests.
5. KMO measure and Bartlett's Test of Sphericity.
6. Exploratory Factor Analysis: Principal Axis Factoring (PAF), Varimax rotation, five factors a priori. Loading criterion: primary loading $\geq .35$ on intended factor (Comrey & Lee, 1992; Stevens, 2002).
7. Pearson inter-trait correlation matrix (Phase 1).
8. Phase 2 cross-validation: internal consistency, split-half reliability, item-total correlations, and discrimination indices. No EFA conducted (N:item = 2.5:1; minimum 10:1 required).

4. RESULTS

4.1 Descriptive Statistics

Table 2 presents trait-level descriptive statistics for both phases on the revised 40-item scale. In Phase 1 (N = 521), mean trait scores ranged from 24.75 (Agreeableness) to 28.02 (Conscientiousness) on the 8–40 scale, all in the mid-range with no floor or ceiling effects. Phase 2 (N = 101) means were comparable (24.28–29.49), indicating score stability across independent samples.

Table 2
Descriptive Statistics — Revised 40-Item Scale, Phase 1 (N = 521) and Phase 2 (N = 101)

Trait	k	P1 N	P1 M	P1 SD	P2 N	P2 M	P2 SD
Conscientiousness	8	521	28.02	6.46	101	29.49	6.49
Openness to Experience	8	521	24.88	6.74	101	25.30	7.50
Neuroticism	8	521	25.18	6.66	101	25.37	7.26

Trait	k	P1 N	P1 M	P1 SD	P2 N	P2 M	P2 SD
Agreeableness	8	521	24.75	6.48	101	25.03	6.77
Extraversion	8	521	25.00	6.36	101	24.28	7.10

Note. k = items per trait. M = Mean trait total. SD = Standard Deviation. Possible range per trait = 8–40. P1 N = 521 after listwise deletion; EFA on N = 520.

4.2 Reliability

Table 3 presents reliability results. Phase 1 Cronbach's alpha ranged from .755 to .794 per trait (full-scale $\alpha = .860$). All Phase 2 alphas improved to the Good range (.827–.894; full-scale $\alpha = .879$). Spearman-Brown corrected split-half coefficients ranged from .847 to .890 (Phase 1) and .874 to .933 (Phase 2). No item upon deletion improved trait alpha by more than .01 in either phase.

Table 3
Internal Consistency and Split-Half Reliability — Phase 1 (N = 521) and Phase 2 (N = 101)

Trait	k	P1 α	P1 SB- ρ	P2 α	P2 SB- ρ	$\Delta\alpha$	P2 Category
Conscientiousness	8	.794	.890	.855	.874	+.061	Good
Openness to Experience	8	.785	.847	.876	.933	+.091	Good
Neuroticism	8	.766	.862	.894	.895	+.128	Good
Agreeableness	8	.770	.865	.834	.894	+.064	Good
Extraversion	8	.755	.855	.827	.915	+.072	Good
Full Scale (40 items)	40	.860	.918	.879	.882	+.019	Good

Note. SB- ρ = Spearman-Brown corrected split-half (odd-even split). $\Delta\alpha = P2 \alpha - P1 \alpha$. Acceptable = .70–.79; Good = .80–.89; Excellent $\geq .90$ (George & Mallery, 2003).

4.3 Discrimination Indices

All 40 items showed statistically significant upper-lower 27% differences (all $p < .001$). Table 4 summarizes discrimination by trait. Phase 1 DI values ranged from .330 to .654; Phase 2 DI values ranged from .250 to .778. The vast majority of items fell in the Good ($\geq .30$) or Excellent ($\geq .40$) range in both phases.

Table 4
Discrimination Index Summary by Trait — Phase 1 (N = 521) and Phase 2 (N = 101)

Trait	n (27%)	P1 Min	P1 Max	P1 Mean	P2 Min	P2 Mean
Conscientiousness	~141	.407	.620	.490	.250	.510
Openness to Experience	~141	.350	.623	.513	.454	.587
Neuroticism	~141	.330	.654	.520	.500	.610
Agreeableness	~141	.404	.625	.492	.444	.530
Extraversion	~141	.387	.562	.473	.472	.540

Note. DI = (Upper 27% M – Lower 27% M) / Score range. Excellent $\geq .40$; Good .30–.39; Fair .20–.29 (Ebel & Frisbie, 1991). All group differences $p < .001$ in both phases. $n = 27\%$ of trait N.

4.4 KMO and Bartlett's Test of Sphericity

Table 5 presents factor analysis prerequisites. Full-scale KMO = .776 (Middling; Kaiser, 1974), all trait KMO values ranged from .722 to .789. All Bartlett's tests significant ($p < .001$). Factor analysis was appropriate.

Table 5
KMO and Bartlett's Test of Sphericity — Phase 1

Scale / Trait	KMO	Category	Bartlett χ^2	df	p
Full Scale (40 items)	.776	Middling	$\approx 10,240$	780	$< .001$
Conscientiousness	.750	Middling	$\approx 2,196$	28	$< .001$
Openness	.789	Middling	$\approx 1,792$	28	$< .001$
Neuroticism	.722	Middling	$\approx 1,876$	28	$< .001$
Agreeableness	.728	Middling	$\approx 1,741$	28	$< .001$
Extraversion	.739	Middling	$\approx 1,658$	28	$< .001$

Note. KMO: $\geq .90$ Marvelous; $\geq .80$ Meritorious; $\geq .70$ Middling; $\geq .60$ Mediocre; $< .60$ Miserable (Kaiser, 1974). EFA method: Principal Axis Factoring, Varimax rotation, 5 factors.

4.5 Exploratory Factor Analysis and Item Classification

EFA was conducted on Phase 1 data ($N = 520$) using Principal Axis Factoring with Varimax rotation, specifying five factors a priori. Table 6 presents variance explained. The five factors accounted for 41.6% of total variance, with eigenvalues ranging from 2.75 to 7.97 — all above the Kaiser criterion of 1.0.

Applying the loading criterion of $\geq .35$ on the intended factor, the 40 retained items are classified as follows:

- 32 items loaded correctly and strongly — primary loading $\geq .35$ on intended factor. These items fully met the structural validity criterion.
- 1 item (Q47, Openness to Experience) loaded on its correct factor (F1) but below the loading threshold ($\lambda = .271 < .35$). It is on the right factor but contributes weakly to its factor definition.
- 7 items showed factor migration — primary loading on a non-intended factor: Q31, Q36, Q41 (Conscientiousness $\rightarrow$ Neuroticism factor); Q3, Q8, Q13 (Neuroticism $\rightarrow$ Openness factor); Q9 (Agreeableness $\rightarrow$ Openness factor, $\lambda = .373$ vs $.292$ on intended Agreeableness).
- Total: $32 + 1 + 7 = 40$ items. Primary factor agreement (items loading on correct factor, regardless of strength) = $33/40 = 82.5\%$. Items meeting both criteria (correct factor AND $\geq .35$) = $32/40 = 80.0\%$.

The 8 flagged items (1 weak-loading + 7 migrating) are retained because they contribute positively to internal consistency and discrimination, but are recommended for wording revision in future iterations.

Table 6
EFA Variance Explained — Phase 1, Revised 40-Item Scale ($N = 520$)

Factor	Trait Label	Eigenvalue	SS Loadings	% Variance	Cumulative %
F1	Openness to Experience	7.97	4.68	9.4	9.4

Factor	Trait Label	Eigenvalue	SS Loadings	% Variance	Cumulative %
F2	Conscientiousness	3.34	4.26	8.5	17.9
F3	Extraversion	3.10	3.84	7.7	25.6
F4	Agreeableness	2.85	3.68	7.4	33.0
F5	Neuroticism	2.75	3.56	7.1	40.1

Note. N = 520. Extraction: Principal Axis Factoring (PAF). Rotation: Varimax (orthogonal). Five factors specified a priori per Big Five theory. Loading criterion $\geq .35$. Item classification: 32 items correct + strong ($\geq .35$ on intended factor); 1 item correct + weak (Q47, $\lambda = .271$); 7 items migrated (wrong primary factor). Total = 40.

4.6 Item Removal: Original 50-Item to Revised 40-Item

The original 50-item scale was first analyzed. Only 27 of 50 items (54%) achieved primary factor agreement. Two items per trait were removed based on systematic and consistent factor migration, yielding a balanced 40-item scale. Removed items: Q4, Q5, Q17, Q18, Q21, Q33, Q34, Q35, Q42, Q46. Following removal, primary factor agreement improved from 54% to 82.5% — a gain of 28.5 percentage points.

4.7 Phase 2 Cross-Validation

Phase 2 (N = 101) cross-validated the reliability of the revised 40-item scale. No EFA was conducted — the 2.5:1 N:item ratio is below the 10:1 minimum for stable factor extraction (McCullum et al., 1999). All five trait alphas improved into the Good range (Table 3). Item-total correlations ranged from .478 to .833 (M = .712); no item fell below .30. Phase 2 discrimination indices ranged from .250 to .778 with all group differences significant at $p < .001$.

4.8 Inter-Trait Correlations

Table 7 shows the Phase 1 inter-trait correlation matrix. Conscientiousness showed near-zero correlations with all other traits ($r = -.018$ to $.108$). The remaining four traits showed moderate positive inter-correlations. The Openness-Neuroticism correlation ($r = .330$) is discussed in Section 5.4.

Table 7
Inter-Trait Pearson Correlation Matrix — Phase 1 (n = 521)

Trait	C	O	N	A	E	M (SD)
Conscientiousness (C)	—	-.018	.031	.061	.108	28.02 (6.46)
Openness (O)	-.018	—	.330**	.268**	.193**	24.88 (6.74)
Neuroticism (N)	.031	.330**	—	.251**	.279**	25.18 (6.66)
Agreeableness (A)	.061	.268**	.251**	—	.272**	24.75 (6.48)
Extraversion (E)	.108	.193**	.279**	.272**	—	25.00 (6.36)

Note. ** $p < .001$. C = Conscientiousness; O = Openness; N = Neuroticism; A = Agreeableness; E = Extraversion. Revised 40-item scale. — = self-correlation.

5. DISCUSSION

The two-phase study gives a clear and honest answer: the revised 40-item questionnaire works well for

Class 10 students in Punjab government schools — with a small number of items that still need attention. We address the key findings in turn.

5.1 Reliability: Consistent Improvement Across Samples

The most important single finding is the consistent alpha improvement from Phase 1 to Phase 2. All five trait alphas moved from the Acceptable to the Good range. Neuroticism showed the largest gain (+.128), reaching $\alpha = .894$. The full-scale alpha of .879 in Phase 2 — across 40 items and five traits in a newly developed scale — is a strong result. This pattern is precisely what a well-functioning scale produces: reliability holds or improves in a second independent sample. The alpha-if-item-deleted analysis confirmed that all 40 items contribute positively in both phases.

5.2 Discrimination: All Items Are Functioning

Every one of the 40 items showed a statistically significant difference between the top and bottom 27% of scorers on its trait (all $p < .001$ in both phases). The average discrimination index exceeded .50 in Phase 2 — in the Excellent range. This means the scale not only measures personality traits but actively and effectively differentiates students at different levels of each trait.

5.3 Factor Structure: Progress and Remaining Challenges

The five-factor structure is clearly present in the data. Moving from 27/50 (54%) correct factor placement in the original scale to 33/40 (82.5%) in the revised scale is a meaningful improvement. Extraversion (8/8) and Openness (7/8) achieved the highest factorial purity. Agreeableness (7/8) was nearly clean. The 32 items meeting both criteria — correct factor and loading $\geq .35$ — represent 80% of the revised scale and perform exactly as the Big Five model predicts.

The 8 flagged items are honest findings, not fatal flaws. Seven items still migrate to non-intended factors, and Q47 loads correctly but weakly. These items are retained because they contribute positively to reliability and discrimination. They are the clearest candidates for wording revision in any future iteration.

5.4 The Conscientiousness-Neuroticism Overlap: A Cultural Perspective

The overlap between Conscientiousness and Neuroticism — three Conscientiousness items loading on the Neuroticism factor — is not simply a measurement problem. In Western normative samples, these two traits are largely orthogonal: being organised and being anxious are treated as unrelated personality dimensions. For Class 10 students in Punjab, however, the experience of being a disciplined, hard-working student may be inseparable from the experience of being under pressure. Board examinations carry significant consequences for stream selection and family expectations; working hard and worrying about performance may co-occur in this population to a degree that simply does not appear in Western samples.

This is an empirically interesting finding that extends the cross-cultural personality literature (Mathur & Gupta, 2020). Future qualitative and longitudinal research into this relationship among Indian adolescents — including how it varies by gender, school setting, and examination context — could contribute meaningfully to both cross-cultural personality theory and educational practice.

5.5 Phase 2 and the Limits of Small-Sample Cross-Validation

Phase 2 ($N = 101$) confirms reliability and item performance. It cannot confirm factor structure — the 2.5:1 N:item ratio is well below the 10:1 minimum for stable EFA (McCullum et al., 1999). An EFA at this sample size would produce mathematically artefactual results in which factor labels re-shuffle freely and do not reflect true underlying structure. For this reason, no Phase 2 EFA was conducted or reported. The Phase 1 five-factor structure ($N = 520$) remains the structural reference. A Phase 3 study

with $N \geq 400$ would enable a full EFA replication and Confirmatory Factor Analysis (CFA) with model fit indices (CFI, TLI, RMSEA, SRMR).

The Phase 2 sample was entirely female (girls' school sections, $n = 101$). While Phase 1 included both boys and girls (56.6% female, 43.4% male), the Phase 2 reliability findings should be interpreted with reference to this female-only sample. Measurement invariance across gender was not tested in either phase — future studies with balanced samples should formally examine whether item functioning and factor structure are equivalent for boys and girls.

5.6 Practical Implications

The validated 40-item scale has immediate practical value for school counselors in Punjab, who can use it to develop evidence-based personality profiles for Class 10 students at the critical juncture of stream selection. High Conscientiousness predicts disciplined follow-through across all streams. High Openness suggests receptiveness to intellectually demanding or creative careers. High Extraversion may indicate aptitude for people-facing roles. High Neuroticism warrants attention to examination-related distress. For researchers, the scale provides a reliable, locally normed measurement tool in a practical 40-item, 20-minute, paper-administered format.

6. CONCLUSION

This study provides the first locally developed and psychometrically validated Big Five Personality Scale for Class 10 students in government schools across Punjab, India. The revised 40-item scale demonstrates good internal consistency ($\alpha = .827-.894$, Phase 2), Good to Excellent item discrimination (all $p < .001$), appropriate factor analytic prerequisites ($KMO = .776$, Bartlett $p < .001$), and a recoverable five-factor structure (33/40 items, 82.5%, on their intended factor; 32/40, 80.0%, meeting the $\geq .35$ loading criterion). Eight items are flagged for wording revision.

The scale is ready for use as a research instrument and counseling tool with this population. A Phase 3 study with $N \geq 400$ is recommended to enable CFA and definitive structural validation. The study also demonstrates that locally grounded psychometric development is both feasible and necessary for Indian adolescent populations, and that the personality experience of Class 10th students in Punjab — including the culturally meaningful Conscientiousness-Neuroticism overlap — deserves investigation through instruments built from within, not borrowed from outside, this community.

REFERENCES

1. Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47(1), i–171.
2. Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38(4), 476–506.
3. Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, 37(4), 319–338.
4. Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319.
5. Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum.
6. Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Psychological Assessment Resources.

7. Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory professional manual. Psychological Assessment Resources.
8. Ebel, R. L., & Frisbie, D. A. (1991). Essentials of educational measurement (5th ed.). Prentice Hall.
9. Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299.
10. George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Allyn & Bacon.
11. Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216–1229.
12. John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). Guilford Press.
13. Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36.
14. Kaur, R., & Verma, S. (2021). Need for culturally responsive personality assessment tools in North Indian secondary schools. *Indian Journal of Applied Psychology*, 58(2), 112–124.
15. Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). Routledge.
16. Lodhi, P. H., Deo, S., & Belhekar, V. M. (2002). The Five-Factor Model of personality in Indian context. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor Model of personality across cultures* (pp. 227–248). Kluwer Academic.
17. MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99.
18. Mathur, S., & Gupta, A. (2020). Big Five personality factors among Indian adolescents: A factor analytic study. *Psychological Studies*, 65(1), 45–58.
19. McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81–90.
20. 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.
21. Nofle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93(1), 116–130.
22. Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
23. Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338.
24. Singh, A., & Bhargava, D. (2015). Adaptation challenges for Western psychological instruments in Indian educational settings. *Journal of Indian Education*, 41(3), 78–92.
25. Srivastava, R., & Misra, G. (2019). Personality assessment in Indian school counselling: A critical review. *Psychology and Developing Societies*, 31(1), 56–81.
26. Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Lawrence Erlbaum.
27. Tupes, E. C., & Christal, R. E. (1961). Recurrent personality factors based on trait ratings (Technical Report ASD-TR-61-97). US Air Force.
28. Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806–838.