Scientific Research Publishing

Psychology, 2026, 17(11), 1264-1279
https://www.scirp.org/journal/psych
ISSN Online: 2152-7199
ISSN Print: 2152-7180

# Nine Labels, Six Signals: Internal Consistency and Exploratory Factor Structure of the 35-Item Multiple Intelligences Profiling Questionnaire III among Undergraduates at a Ghanaian University

**Aba Arthur-Sarfo[1], Juliana Dontoh[2], Regine Kwaw Nsiah[3], Prince Yaw Adofo[4], Inuusah Mahama[5*]**

[1]Department of Education, Ola College of Education, Cape Coast, Ghana
[2]Department of Education Studies, Accra College of Education, Accra, Ghana
[3]Department of Education, Enchi College of Education, Enchi, Ghana
[4]School of Arts and Social Sciences Education, Ghana Baptist University College, Kumasi, Ghana
[5]Department of Counselling Psychology, University of Education, Winneba, Ghana
Email: *imahama@uew.edu.gh

**How to cite this paper:** Arthur-Sarfo, A., Dontoh, J., Nsiah, R. K., Adofo, P. Y., & Mahama, I. (2026). Nine Labels, Six Signals: Internal Consistency and Exploratory Factor Structure of the 35-Item Multiple Intelligences Profiling Questionnaire III among Undergraduates at a Ghanaian University. *Psychology, 17,* 1264-1279.
https://doi.org/10.4236/psych.2026.1711062

**Received:** October 6, 2026
**Accepted:** November 7, 2026
**Published:** November 10, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

## Abstract

Background: The 35-item Multiple Intelligences Profiling Questionnaire III (MIPQ III) provides scores for nine self-perceived domains, but their internal consistency and latent structure require contextual validation. Method: Responses from 613 volunteers at a Ghanaian university were analysed. After excluding two complete invariant records, 537 cases with all 35 item responses were available for principal-axis exploratory factor analysis with varimax rotation. Parallel analysis used 1000 random permutations of the observed item distributions and compared full-correlation eigenvalues with their simulated 95th percentiles. Results: Internal consistency varied from $\alpha = 0.373$ (linguistic) to $\alpha = 0.835$ (musical). The Kaiser-Meyer-Olkin index was 0.896. Full-correlation parallel analysis retained six factors. A reflective-environmental cluster combined intrapersonal, spiritual and environmental items, while musical, mathematical, interpersonal and practical-motor item clusters were more recognisable; a sixth factor contrasted two school-subject items. Two items showed cross-loadings of at least 0.30 on multiple factors. The six-factor retention decision persisted after limited missing-value imputation, although reduced common-factor parallel analysis favoured seven factors. In a forced nine-factor model, unconstrained matching showed that intrapersonal and spiritual items loaded strongly on a factor already allocated to the environmental domain.

Conclusion: The nine domain scores exhibited unequal reliability and insufficient structural separation to warrant interpretation as nine independently measured abilities. The MIPQ III may be useful for exploratory self-perception research, but performance validation and cross-sample replication are needed.

## Keywords

Multiple Intelligences, MIPQ III, Psychometric Evaluation, Internal Consistency, Exploratory Factor Analysis, Ghana

## 1. Introduction

The theory of multiple intelligences proposed a broad account of human capability and has influenced how educators discuss students' diverse strengths (Gardner, 1983, 1999). Yet the educational appeal of a theory does not establish that a short questionnaire can separate the proposed domains with sufficient precision. Gardner and Moran (2006) distinguished theoretical propositions from the validity of particular operational measures, while empirical critiques have questioned whether the intelligences are psychometrically distinct (Visser et al., 2006; Waterhouse, 2006). Consequently, instruments that assign individuals to intelligence profiles require direct examination of score consistency and latent structure.

The multiple-intelligences framework continues to shape educational discussions of learner diversity, broad opportunity and the interpretation of strengths (Gardner, 2024a, 2024b). Research on teachers' use of multiple-intelligences approaches illustrates its pedagogical appeal (Cianci, 2024), while work on learner intelligence profiles illustrates continuing interest in profile-based description (Corraya et al., 2025). These applications should not, however, be confused with evidence that questionnaire scores reproduce distinct latent dimensions. A long-standing distinction in the study of psychological attributes is between a construct and the particular indicators selected to assess it (Eysenck, 2026).

Earlier work has examined the measurement of multiple intelligences rather than assuming that the proposed domains are automatically separable (Tirri et al., 2013). Spirituality and giftedness also form a specialised conceptual literature (Tirri, 2023; Çiftçi, 2025), and recent assessment work in Ghana and Botswana considers spiritual intelligence in particular educational settings (Ntumi et al., 2026). These related publications provide conceptual and contextual reasons to examine the MIPQ III, not independent validation of a nine-factor structure among Ghanaian undergraduates.

The Multiple Intelligences Profiling Questionnaire III (MIPQ III) operationalises nine domains using 35 ordered self-report items: linguistic, logical-mathematical, spatial, bodily-kinesthetic, musical, interpersonal, intrapersonal, spiritual and environmental (Tirri & Nokelainen, 2008, 2011). Each domain is represented by only three or four statements. Brief scales can be practical in educational sur-

veys, but they pose psychometric challenges: high interitem agreement may coexist with strong overlap between intended constructs, whereas low agreement may undermine individual-level interpretation (Cronbach, 1951; McNeish, 2018). Moreover, endorsing a statement about a perceived strength does not necessarily demonstrate proficiency in that domain.

Self-reported proficiency may reflect perceived ability, motivational investment and social identity rather than tested performance. Personality and motivation research shows why these elements should be conceptually distinguished (McAdams & Dunlop, 2022; Scheffer & Heckhausen, 2025). Empirical work comparing children's self-reports with teacher ratings in a different assessment area likewise demonstrates the importance of considering informant and method (Piombo et al., 2026).

Reliability also requires a proportionate interpretation. Cronbach's (1951) coefficient is informative about item covariance, but alternative reliability coefficients and the assumptions attached to alpha have been debated (Revelle & Zinbarg, 2009; McNeish, 2018). Evidence concerning the changing internal consistency of domain-specific knowledge assessments further cautions against treating a single coefficient as a complete description of measurement quality (Edelsbrunner et al., 2025). Reviews of questionnaire development recommend examining content validity, score distributions and the match between items and intended constructs alongside coefficient estimates (Hossan et al., 2025; Zumitzavan, 2026). Ordered-category items additionally warrant attention to ordinal reliability (Gadermann et al., 2012).

These concerns are relevant in Ghanaian higher education, where students may have varied language backgrounds, educational experiences and opportunities to demonstrate abilities. The present study therefore evaluated the instrument as a measure of self-perceived tendencies rather than treating its domain labels as established representations of independent cognitive capacities. It addressed three questions: How internally consistent are the nine keyed scores? How many exploratory dimensions are retained using parallel analysis? To what extent do the observed loading patterns distinguish or combine the nine intended domains? The study additionally tested whether factor retention depended on missing-response decisions or the correlation-matrix convention used in parallel analysis.

## 2. Method

### 2.1. Design, Recruitment and Participants

A cross-sectional psychometric study was conducted using responses submitted through a voluntary online questionnaire. The dataset contains 613 undergraduate records from one university in Ghana. The sample is a nonprobability volunteer sample; the source documentation does not specify the invitation channel, recruitment period, number of students eligible or invited, or the academic programmes represented. Accordingly, no defensible response rate can be calculated and representativeness cannot be presumed. These characteristics are limitations

of the existing recruitment documentation, not details inferred from the questionnaire responses.

Of the 613 records, 419 (68.35%) were marked male and 183 (29.85%) female; gender was unrecorded for 11 (1.79%). Students reported Level 300 (n = 366, 59.71%), Level 200 (n = 182, 29.69%), Level 400 (n = 54, 8.81%) or Level 100 (n = 2, 0.33%); year level was missing for nine (1.47%). Missing gender entries were not classified as a nonbinary category.

## 2.2. Instrument and Scoring

The 35-item MIPQ III (Tirri & Nokelainen, 2008, 2011) uses five ordered response categories coded 1 - 5. Items 1 - 4 measure linguistic self-perceptions; items 5 - 8, logical-mathematical; items 9 - 12, spatial; items 13 - 16, bodily-kinesthetic; items 17 - 20, musical; items 21 - 24, interpersonal; items 25 - 28, intrapersonal; items 29 - 32, spiritual; and items 33 - 35, environmental self-perceptions. Following the original scoring rule, each domain score was the mean of its observed items if at least 75% were answered. For four-item domains at least three responses were required; the environmental score required all three. Higher scores represent stronger self-endorsement, not an objectively demonstrated level of ability.

## 2.3. Data Screening, Missing Responses and Analytic Samples

All observed item values were within the allowed 1 - 5 range. The data matrix comprised 613 × 35 = 21,455 possible item responses, of which 217 (1.01%) were missing. In total, 539 respondents had complete responses to all 35 items; 74 respondents had at least one missing item. Two of the 539 fully complete response vectors were invariant across all 35 items: one respondent marked 1 throughout and another marked 5 throughout. These were excluded from the primary analyses because a constant response across heterogeneous domains could indicate indiscriminate response behaviour; it is not, however, proof of poor-quality data. The primary factor analytic dataset therefore contained 537 complete, noninvariant records (87.60% of all cases).

Reliability coefficients were estimated using all respondents excluding the two complete invariant cases, with complete observations required within the particular domain; descriptive scores used the stated 75% rule. Thus, domain-level sample sizes differed. Complete-case analysis was chosen for the primary correlation matrix and factor solutions to ensure a common 537-person basis for all 35 item correlations and direct comparability with the original analysis. This choice assumes that restricting the sample does not materially alter the correlation structure and can introduce bias when item missingness is not completely random. Two sensitivity procedures were undertaken: reincluding the two invariant complete cases and single median imputation for respondents missing no more than two item responses (n = 603, excluding the invariant cases). A further check included up to four missing items per respondent (n = 606). Median imputation was used strictly as a robustness check, not as a preferred method for estimating pop-

ulation parameters.

## 2.4. Statistical Analysis and Reproducibility

Analyses were rerun in Python 3.13.5 with NumPy 2.3.5, SciPy 1.17.0, and statsmodels 0.14.6. Reliability was assessed using raw-score Cronbach's alpha, mean Pearson interitem correlations and the range of corrected item-total correlations. Domain means, standard deviations and ceiling proportions were calculated for all eligible noninvariant records. Factorability was examined with the Kaiser-Meyer-Olkin (KMO) index and Bartlett's test of sphericity.

Factor analysis and principal-component analysis answer related but non-identical questions; this distinction matters when interpreting eigenvalues and latent factors (Allee et al., 2022; Garson, 2022). Methodological discussions distinguish parallel-analysis conventions and stress the consequences of the correlation-matrix definition (Holubova, 2023; Mardia et al., 2024). Recommendations for exploratory factor analysis also emphasise transparent extraction, retention and rotation decisions and substantive examination of loading patterns rather than an eigenvalue rule alone (Fabrigar et al., 1999; Sürücü et al., 2024). These principles motivated the primary and reduced-matrix analyses reported below.

For the primary retention analysis, observed eigenvalues were obtained from the full 35 × 35 Pearson item correlation matrix of the 537 complete, noninvariant records. Parallel analysis used 1000 Monte Carlo replicates with NumPy's default pseudo-random generator (seed 20261016); in each replicate, the observed responses in each item column were independently permuted without replacement, preserving that item's empirical marginal distribution while removing between-item associations. Each observed eigenvalue was compared with the corresponding 95th-percentile random eigenvalue. The retained factor count was the number of consecutive ranks beginning at rank 1 whose observed eigenvalues exceeded the random benchmark. For clarity, the full-correlation eigenvalues are principal-component-style retention statistics, although the factor loadings were subsequently estimated using common-factor extraction.

Principal-axis factoring, with squared multiple correlations as initial communality estimates, was used to extract six factors; orthogonal varimax rotation was applied for direct comparison with the intended simple structure. Salient pattern magnitudes were interpreted at $|\lambda| \geq 0.40$, and potential cross-loadings at $|\lambda| \geq 0.30$ on at least two factors; these thresholds were descriptive aids rather than significance tests. A separately fitted nine-factor principal-axis solution with varimax rotation examined the theory-constrained key. For each keyed domain, the mean absolute loading was computed on every factor and its strongest factor identified without restriction. This unrestricted result was contrasted with a global one-to-one allocation, calculated using the Hungarian assignment algorithm to maximise the summed mean absolute loadings across the nine domains. The number of domain items with their largest absolute loading on each candidate factor was also counted. Factor signs are arbitrary and were oriented in **Table 3** to aid interpre-

tation.

As an additional method sensitivity check, parallel analysis was repeated using reduced Pearson correlation matrices with squared multiple correlations replacing the diagonal, both for the observed data and for each simulated matrix. The identical sequential 95th-percentile decision rule was applied. This reduced-matrix procedure targets common-factor rather than full-matrix eigenvalue retention and should not be treated as interchangeable with the primary approach (Horn, 1965; Hayton et al., 2004). Independent normally distributed Monte Carlo data were also used as an auxiliary check of full-matrix retention.

## 3. Results

### 3.1. Domain Descriptives and Internal Consistency

Domain descriptives are given in Table 1. Environmental self-perceptions were highest (M = 4.28, SD = 0.91), followed by intrapersonal (M = 4.05, SD = 0.91) and spiritual (M = 3.79, SD = 0.82) scores. The environmental scale showed a marked ceiling: 37.79% of scorable respondents obtained the maximum possible mean. This restricts differentiation among respondents at the upper end of the scale.

Internal consistency was uneven. Musical (α = 0.835), intrapersonal (α = 0.810) and environmental (α = 0.802) scores had the largest coefficients. Linguistic (α = 0.373) and spatial (α = 0.418) scores were notably weak, with mean interitem correlations of 0.13 and 0.15. The remaining domains had intermediate coefficients (α = 0.591 - 0.658). These results distinguish reliability within each proposed domain from the separate question of whether domains can be discriminated factorially.

Table 1. **MIPQ III** descriptives and reliability by keyed domain.

| Domain | k | n | M | SD | α | Mean r | Ceiling % |
|---|---|---|---|---|---|---|---|
| Linguistic | 4 | 609 | 3.553 | 0.81 | 0.373 | 0.13 | 5.91 |
| Logical-mathematical | 4 | 606 | 3.287 | 0.86 | 0.613 | 0.29 | 3.63 |
| Spatial | 4 | 607 | 3.327 | 0.731 | 0.418 | 0.15 | 2.47 |
| Bodily-kinesthetic | 4 | 606 | 3.497 | 0.917 | 0.648 | 0.31 | 7.43 |
| Musical | 4 | 605 | 3.281 | 1.113 | 0.835 | 0.56 | 10.25 |
| Interpersonal | 4 | 606 | 3.577 | 0.881 | 0.591 | 0.27 | 9.74 |
| Intrapersonal | 4 | 607 | 4.053 | 0.908 | 0.810 | 0.52 | 17.96 |
| Spiritual | 4 | 606 | 3.794 | 0.818 | 0.658 | 0.33 | 9.08 |
| Environmental | 3 | 598 | 4.283 | 0.914 | 0.802 | 0.58 | 37.79 |

Note. All scores exclude the two complete invariant responders. n denotes the number meeting the 75% item-completeness rule; Cronbach's alpha used cases complete within each domain (n = 591 - 601 for four-item domains; 598 for environmental). Mean r = mean interitem correlation. Ceiling = percentage with the maximum possible domain score.

### 3.2. Factorability and Parallel Analysis

Sampling adequacy was high (KMO = 0.896). Bartlett's test rejected the hypothesis of an identity item correlation matrix, $\chi^2(595) = 6541.16$, $p < 0.001$. In the primary full-correlation parallel analysis, the first six observed eigenvalues (8.548, 2.661, 2.018, 1.481, 1.435 and 1.367) exceeded their respective simulated 95th-percentile values. The seventh observed eigenvalue (1.187) was below its random benchmark (1.289). On this explicitly defined retention procedure, six factors were indicated (Table 2). Numerical ranks in the table do not denote linguistic, musical or any other prespecified intelligence domain.

Table 2. Full-correlation parallel analysis by eigenvalue rank.

| Rank | Observed | Random mean | Random 95th | Decision |
|---|---|---|---|---|
| 1 | 8.548 | 1.514 | 1.574 | Retain |
| 2 | 2.661 | 1.452 | 1.496 | Retain |
| 3 | 2.018 | 1.404 | 1.447 | Retain |
| 4 | 1.481 | 1.362 | 1.397 | Retain |
| 5 | 1.435 | 1.325 | 1.359 | Retain |
| 6 | 1.367 | 1.291 | 1.320 | Retain |
| 7 | 1.187 | 1.261 | 1.289 | Do not retain |
| 8 | 1.070 | 1.230 | 1.257 | Do not retain |
| 9 | 0.974 | 1.201 | 1.227 | Do not retain |

Note. Observed eigenvalues are from the complete-case Pearson correlation matrix (n = 537). Monte Carlo benchmark: 1000 independently permuted item-column datasets, seed 20261016; retention uses the 95th percentile at each rank. These are eigenvalue ranks, not intelligence-domain labels.

### 3.3. Interpretation of the Six-Factor Structure

Principal-axis factoring with varimax rotation yielded six item dimensions whose squared loading sums accounted for 41.21% of total standardised item variance. The most extensive factor combined intrapersonal items 25 - 28, spiritual items 29 - 31 and environmental items 33 - 35, plus selected items from other keyed domains. The musical items 17 - 20 formed a comparatively coherent second factor. A third factor was principally associated with logical-mathematical items 6 - 8 and spatial item 10; a fourth was dominated by interpersonal items 21 and 23; and a fifth captured practical/manual items 14 - 16. The sixth was not a convincing additional intelligence scale: it primarily contrasted the school-language/social-studies item 2 with the mathematics/science item 5. Thus, the retained six factors were neither equally well defined nor straightforward equivalents of six proposed intelligences.

Two items showed cross-loadings of at least 0.30 on two factors: item 8 (logical

reasoning and games) loaded 0.543 on factor 3 and 0.330 on factor 1; item 15 (showing others how to do something) loaded 0.488 on factor 1 and 0.444 on factor 5. Five items (1, 3, 11, 13 and 22) had no loading with magnitude reaching 0.30. These results indicate that the original items do not uniformly define distinct simple-structure constructs. Table 3 displays the complete six-factor solution rather than merely reporting the factor count.

Table 3. Varimax-rotated loadings for the retained six-factor principal-axis solution.

| Item | Key | F1 | F2 | F3 | F4 | F5 | F6 | $h^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | Ling. | 0.192 | 0.080 | 0.130 | −0.041 | −0.042 | 0.060 | 0.067 |
| 2 | Ling. | 0.270 | 0.082 | 0.017 | 0.055 | 0.006 | **+0.628** | 0.477 |
| 3 | Ling. | 0.151 | 0.127 | 0.225 | 0.136 | 0.031 | −0.033 | 0.110 |
| 4 | Ling. | 0.312 | 0.091 | 0.249 | 0.104 | 0.040 | 0.165 | 0.207 |
| 5 | Logic. | −0.048 | 0.016 | 0.250 | −0.014 | 0.131 | **−0.531** | 0.365 |
| 6 | Logic. | 0.299 | −0.005 | **+0.530** | 0.054 | 0.083 | −0.174 | 0.411 |
| 7 | Logic. | 0.134 | 0.052 | **+0.526** | 0.018 | 0.080 | −0.259 | 0.371 |
| 8 | Logic. | 0.330 | 0.095 | **+0.543** | 0.036 | 0.143 | 0.069 | 0.440 |
| 9 | Spatial | −0.020 | 0.155 | 0.312 | 0.091 | 0.093 | 0.064 | 0.142 |
| 10 | Spatial | 0.125 | 0.102 | **+0.441** | 0.058 | 0.092 | 0.117 | 0.246 |
| 11 | Spatial | 0.138 | 0.126 | 0.227 | 0.137 | 0.150 | 0.170 | 0.157 |
| 12 | Spatial | **+0.590** | 0.148 | 0.173 | 0.061 | 0.095 | 0.098 | 0.422 |
| 13 | Motor | 0.272 | 0.135 | 0.117 | 0.083 | 0.190 | 0.094 | 0.158 |
| 14 | Motor | 0.150 | 0.132 | 0.011 | 0.102 | **+0.710** | −0.053 | 0.557 |
| 15 | Motor | **+0.488** | 0.182 | 0.136 | 0.066 | **+0.444** | 0.086 | 0.498 |
| 16 | Motor | 0.019 | 0.174 | 0.116 | 0.034 | **+0.713** | −0.020 | 0.554 |
| 17 | Music | 0.150 | **+0.635** | 0.097 | 0.079 | 0.086 | 0.083 | 0.456 |
| 18 | Music | 0.123 | **+0.824** | 0.023 | 0.062 | 0.085 | 0.004 | 0.705 |
| 19 | Music | 0.066 | **+0.779** | 0.015 | 0.053 | 0.114 | 0.027 | 0.628 |
| 20 | Music | 0.168 | **+0.696** | 0.029 | 0.087 | 0.059 | −0.025 | 0.525 |
| 21 | Inter. | 0.213 | 0.197 | 0.037 | **+0.570** | 0.028 | −0.044 | 0.414 |
| 22 | Inter. | 0.124 | 0.087 | 0.029 | 0.141 | −0.018 | 0.079 | 0.050 |
| 23 | Inter. | 0.137 | 0.079 | 0.025 | **+0.874** | 0.073 | 0.037 | 0.796 |
| 24 | Inter. | **+0.616** | 0.092 | 0.136 | 0.214 | 0.058 | 0.016 | 0.456 |
| 25 | Intra. | **+0.749** | 0.137 | 0.157 | 0.052 | 0.021 | −0.020 | 0.608 |

**Continued**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 26 | Intra. | **+0.693** | 0.118 | 0.142 | 0.001 | 0.022 | 0.057 | 0.518 |
| 27 | Intra. | **+0.779** | 0.048 | 0.033 | 0.036 | 0.026 | 0.049 | 0.614 |
| 28 | Intra. | **+0.475** | 0.139 | 0.268 | 0.240 | 0.051 | 0.022 | 0.377 |
| 29 | Spirit. | **+0.450** | 0.143 | 0.063 | 0.180 | −0.005 | −0.003 | 0.259 |
| 30 | Spirit. | **+0.561** | 0.062 | 0.049 | 0.064 | 0.029 | 0.114 | 0.339 |
| 31 | Spirit. | **+0.724** | 0.120 | 0.061 | 0.088 | −0.020 | −0.017 | 0.551 |
| 32 | Spirit. | 0.356 | 0.117 | 0.112 | 0.255 | −0.006 | 0.104 | 0.228 |
| 33 | Environ. | **+0.795** | 0.073 | 0.059 | 0.063 | 0.092 | 0.038 | 0.655 |
| 34 | Environ. | **+0.735** | 0.028 | 0.002 | 0.101 | 0.096 | 0.145 | 0.582 |
| 35 | Environ. | **+0.665** | 0.088 | −0.020 | 0.156 | 0.077 | −0.008 | 0.480 |

Note. F1 = broad reflective-environmental cluster; F2 = musical; F3 = mathematical/selected spatial; F4 = interpersonal; F5 = manual/practical; F6 = mixed school-subject contrast. $h^2$ = extracted communality. Signs are arbitrary; factor 5 and factor 6 may contain negative and positive item associations. Magnitudes at or above 0.40 are emphasised in the table text; cross-loadings were screened at 0.30.

### 3.4. Forced Nine-Factor Structure: Unrestricted versus Unique Matches

For comparison with the nine-domain scoring key, a forced nine-factor principal-axis solution accounted for 45.91% of total standardised item variance. Musical and environmental domains were coherent on their respective factors (all four musical items and all three environmental items primarily loaded on their strongest factors). However, Table 4 shows that a unique domain-to-factor assignment substantially underestimated the shared loading structure of intrapersonal and spiritual items. Both domains had their best unrestricted match on factor 1, also preferred by environmental items. Intrapersonal mean absolute loading was 0.682 on factor 1 (four of four primary items), but only 0.166 after unique matching forced it onto factor 9 (zero primary items). The corresponding values for spiritual items were 0.518 on factor 1 (four of four) versus 0.157 on assigned factor 7 (zero of four). These low assigned matches therefore arose mainly from the constraint against sharing a factor, not from an absence of coherent item covariation.

Table 4. Strongest unrestricted and uniquely assigned factors in the nine-factor solution.

| Keyed domain | Best F | Best \|λ\| | Primary | Assigned F | Assigned \|λ\| | Primary | Difference |
|---|---|---|---|---|---|---|---|
| Linguistic | 1 | 0.223 | 2/4 | 6 | 0.219 | 1/4 | 0.004 |
| Logical-mathematical | 3 | 0.478 | 3/4 | 3 | 0.478 | 3/4 | 0.000 |
| Spatial | 3 | 0.251 | 2/4 | 8 | 0.157 | 0/4 | 0.094 |
| Bodily-kinesthetic | 5 | 0.510 | 2/4 | 5 | 0.510 | 2/4 | 0.000 |

**Continued**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Musical | 2 | 0.734 | 4/4 | 2 | 0.734 | 4/4 | 0.000 |
| Interpersonal | 4 | 0.469 | 2/4 | 4 | 0.469 | 2/4 | 0.000 |
| Intrapersonal | 1 | 0.682 | 4/4 | 9 | 0.166 | 0/4 | 0.516 |
| Spiritual | 1 | 0.518 | 4/4 | 7 | 0.157 | 0/4 | 0.361 |
| Environmental | 1 | 0.725 | 3/3 | 1 | 0.725 | 3/3 | 0.000 |

Note. Best F identifies the factor with highest mean absolute loading for that domain without any uniqueness restriction. Assigned F is the factor selected by a global one-to-one matching algorithm. Primary counts reflect the number of domain items whose highest absolute loading is on that factor. The difference is Best |λ| minus Assigned |λ|.

Spatial items were comparatively dispersed (best unrestricted mean absolute loading = 0.251 on factor 3, with two primary items); the linguistic best-factor mean was also low (0.223). In contrast, the musical, environmental, logical-mathematical, bodily-kinesthetic and interpersonal domains retained the same strongest factor under constrained and unconstrained allocation. The findings challenge independent recovery of all nine named constructs without implying that every nominal domain is internally incoherent.

### 3.5. Missing-Data and Parallel-Analysis Sensitivity

The primary six-factor retention decision was unchanged when both invariant records were retained in the complete-case analysis (n = 539), when up to two missing item responses per person were replaced with the observed item medians (n = 603), or when up to four missing responses were so replaced (n = 606). It was also unchanged when the null datasets were generated from independent standard-normal variables rather than permutations of each observed item column. Table 5 summarises these comparisons. However, reduced-matrix parallel analysis with squared multiple correlations on the diagonal retained seven sequential factors in all four sample specifications. The narrow eighth-rank gap in this reduced-matrix

Table 5. Sensitivity of sequential 95th-percentile parallel-analysis retention.

| Analytic specification | n | Full R | Reduced R | Rank-7 gap |
|---|---|---|---|---|
| Complete, noninvariant (primary) | 537 | 6 | 7 | −0.102 |
| All complete records | 539 | 6 | 7 | −0.113 |
| Median imputation: ≤2 missing | 603 | 6 | 7 | −0.139 |
| Median imputation: ≤4 missing | 606 | 6 | 7 | −0.139 |

Note. Full R = full Pearson correlation matrix; Reduced R = observed and random matrices with squared multiple correlations as diagonal estimates. Rank-7 gap = observed minus 95th-percentile random eigenvalue at rank 7 under the full-matrix method. The primary finding is stable to the stated exclusion/imputation changes, but changes from six to seven under reduced-matrix retention. Median imputation is sensitivity analysis only.

procedure underscores that dimension count is partly method dependent. Hence, "six factors" characterises the chosen primary full-correlation retention specification rather than a universally established latent structure.

## 4. Discussion

The study provides a differentiated appraisal of the MIPQ III among the surveyed undergraduates. Although nine scale scores are stipulated by the questionnaire, the evidence does not support treating all nine as internally reliable and factorially independent. Instead, musical, intrapersonal and environmental composites exhibited comparatively high internal consistency, linguistic and spatial composites had low reliability, and the primary factor-retention procedure identified six rather than nine dimensions. The full loading matrix is essential to understanding this result: the first factor included intrapersonal, spiritual and environmental items, whereas a sixth factor reflected a narrow contrast between two school-subject statements rather than a coherent new intelligence.

The unconstrained nine-factor comparisons explain an apparent inconsistency in the original report. Intrapersonal items were reliable within their assigned scale but appeared poorly recovered when a one-to-one matching algorithm forced them away from the factor also preferred by environmental and spiritual items. The direct loading analysis shows that intrapersonal and spiritual items shared a substantial factor with environmental content. Accordingly, reliability is not sufficient evidence of discriminant validity, and a weak forced assignment must not be conflated with a lack of shared structure (Cronbach, 1951; McNeish, 2018; Revelle & Zinbarg, 2009).

The musical domain offered the clearest case of a relatively distinctive four-item self-perception cluster. The environmental items were similarly coherent but showed a pronounced ceiling, limiting precision among highly endorsing respondents. The linguistic and spatial scales are of particular concern, because their weak average interitem associations and diffuse loading patterns compromise score precision. These discrepancies call for item-level content review, cognitive interviews, broader initial item pools, and testing whether differently worded items represent perceived capability, personal interest or opportunities for learning rather than a single latent ability.

The factor-count sensitivity analysis adds an important qualification to the title's six-factor finding. Independent permutation-based parallel analysis of the full correlation matrix repeatedly retained six factors under reasonable data-exclusion choices. Nevertheless, an otherwise comparable reduced common-factor matrix procedure retained seven. Parallel analysis must therefore be identified by matrix convention, simulation method and threshold rather than cited simply as a universal rule (Horn, 1965; Hayton et al., 2004). In addition, treating ordered 1 - 5 responses with Pearson correlations is an approximation; future work should examine ordinal/polychoric analyses, factor stability, and alternative oblique rotations that allow genuine domain associations (Gadermann et al., 2012; Fabrigar et

al., 1999).

External validation should be designed around the inference that an instrument is intended to support. Measurement discussions in other applied disciplines make the same general point: a useful instrument must be matched to its target construct and decision context (Kyle et al., 2025). Studies seeking behavioural indicators beyond self-report, including work on critical-thinking assessment, offer examples of a different measurement strategy (Li et al., 2026). By contrast, exploratory psychophysiological work associated with multiple-intelligences concepts uses a substantially different mode of measurement (Minkin et al., 2024). Neither line of research establishes convergent validity for the 35-item MIPQ III in this sample.

The importance of keeping the assessment target explicit is also evident in research evaluating personality-like outputs from large language models (Serapio-García et al., 2023): properties of a measurement procedure and its respondent cannot simply be transferred to another population or construct. That study does not provide evidence about human intelligence profiles. For the present questionnaire, a defensible next step is to compare item scores with appropriately chosen human performance assessments, observer reports and prospective educational criteria before interpreting them as indicators of ability.

### 4.1. Implications for Education and Assessment

The questionnaire can still serve as a structured prompt for discussing perceived interests, learning experiences and strengths. However, self-ratings should not be interpreted as performance tests, nor should students be assigned to curriculum tracks, denied opportunities or taught exclusively through a presumed intelligence type on the basis of these scores. The unequal reliability across domains means that even descriptive group comparisons should be reported scale by scale, with attention to uncertainty. Educational interpretations would be stronger if supported by independent performance tasks, other-informant ratings and evidence that domain scores predict meaningful outcomes beyond general self-confidence or enjoyment.

### 4.2. Study Limitations

First, the study relied on voluntary participation at one Ghanaian university. Recruitment invitations, eligible population counts and academic-programme representation were not documented in the analytic files. The response rate cannot be calculated, and selection bias limits the representativeness and generalisability of the estimates. Second, the MIPQ III relies entirely on students' self-appraisals, which may reflect interests, response styles, educational opportunity and social desirability as well as perceived competence; no external ability or performance criteria were administered. Third, factor extraction used 537 of the 613 available respondents because of missing responses and two complete invariant response patterns. Although six-factor retention persisted in specified imputation and in-

clusion checks, these are not a substitute for investigating the missingness mechanism or using model-based methods for incomplete ordinal data.

Fourth, the six factors were exploratory rather than confirmed, and several factors contained few well-defined items or substantial content overlap. The sample was not divided into independent exploratory and confirmatory subsamples; confirmatory factor analysis, test-retest reliability, cross-institutional measurement invariance and longitudinal or performance-based validation were not undertaken. Fifth, the primary six-factor conclusion was sensitive to the full-correlation versus reduced-matrix convention of parallel analysis, and an orthogonal varimax solution may obscure associations between latent dimensions. Together, these limitations preclude treating either six or nine factors as a final validated model of distinct intelligences.

## 5. Conclusion

The MIPQ III produced an uneven nine-score profile in this Ghanaian undergraduate sample. Reproducible full-correlation parallel analysis indicated six dimensions under the prespecified retention criterion, but common-factor parallel analysis indicated seven. More importantly, the six-factor loading matrix and unrestricted nine-factor comparisons revealed both distinct item clusters and marked overlap, especially among intrapersonal, spiritual and environmental self-perceptions. The strongest justified conclusion is about the present self-report instrument, not the number of human intelligences: its nine scores should not be assumed to measure nine separable abilities. Further item refinement, independently replicated factor analysis and performance-based validation are needed before consequential individual use.

## Ethics Approval and Informed Consent

Ethical approval was reported from the Department of Counselling Psychology, University of Education, Winneba (reference CPSY/UEW/EDU/09-25/69). Participation was voluntary and respondents provided informed consent. These details are retained from the study documentation supplied by the authors.

## Data Availability

The de-identified data analysed in this study may be requested from the corresponding author, subject to applicable ethical and institutional restrictions. The analytic specifications (factor extraction, randomisation procedure and random seed) are described to support reproduction.

## Acknowledgements and Use of AI Tools

The authors used ChatGPT (OpenAI) for assistance with language editing, manuscript organisation and development of reproducible analysis code. The authors reviewed and take responsibility for the data analysis, interpretations, references and final content.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References


Allee, K. D., Do, C., & Raymundo, F. G. (2022). Principal Component Analysis and Factor Analysis in Accounting Research. *Journal of Financial Reporting, 7,* 1-39. https://doi.org/10.2308/jfr-2021-005

Cianci, M. G. M. (2024). *The Experiences of University Faculty with Integrating Multiple Intelligences into Their Teaching Strategies*. Ph.D. Thesis, Capella University.

Çiftçi, M. (2025). Spirituality and Giftedness: A Brief Review. *International Journal of Children's Spirituality, 30,* 231-240. https://doi.org/10.1080/1364436x.2025.2552128

Corraya, S., Al Mamun, S., & Kaiser, M. S. (2025). Factors Associated with the Adult Learners' Multiple Intelligence Profile. In M. S. Kaiser, A. Bandyopadhyay, M. Mahmud, & K. Ray (Eds.), *Lecture Notes in Networks and Systems* (pp. 413-430). Springer. https://doi.org/10.1007/978-981-95-1069-6_29

Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika, 16,* 297-334. https://doi.org/10.1007/bf02310555

Edelsbrunner, P. A., Simonsmeier, B. A., & Schneider, M. (2025). The Cronbach's Alpha of Domain-Specific Knowledge Tests before and after Learning: A Meta-Analysis of Published Studies. *Educational Psychology Review, 37,* Article No. 4. https://doi.org/10.1007/s10648-024-09982-y

Eysenck, H. J. (2026). *The Scientific Study of Personality*. Routledge. https://doi.org/10.4324/9781003794240

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,* 272-299. https://doi.org/10.1037/1082-989x.4.3.272

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating Ordinal Reliability for Likert-Type and Ordinal Item Response Data: A Conceptual, Empirical, and Practical Guide. *Practical Assessment, Research & Evaluation, 17,* 1-13.

Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.

Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*. Basic Books.

Gardner, H. (2024a). The Essential Howard Gardner on Education. Teachers College Press.

Gardner, H. (2024b). *The Essential Howard Gardner on Mind*. Teachers College Press.

Gardner, H., & Moran, S. (2006). The Science of Multiple Intelligences Theory: A Response to Lynn Waterhouse. *Educational Psychologist, 41,* 227-232. https://doi.org/10.1207/s15326985ep4104_2

Garson, G. D. (2022). *Factor Analysis and Dimension Reduction in R: A Social Scientist's Toolkit*. Routledge. https://doi.org/10.4324/9781003279693

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *Organizational Research Methods, 7,* 191-205. https://doi.org/10.1177/1094428104263675

Holubova, H. (2023). A Comparative Analysis of the Principal Component Method and Parallel Analysis in Working with Official Statistical Data. *Statistics in Transition New Series, 24,* 199-212. https://doi.org/10.59170/stattrans-2023-011

Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis.

*Psychometrika, 30,* 179-185. https://doi.org/10.1007/bf02289447

Hossan, D., Wolfs, B., & Petkovic, M. (2025). Questionnaire Validity and Reliability: A Review with Practical Guidelines. *Journal of Entrepreneurship, Business and Economics, 13,* 135-186.

Kyle, G., Landon, A. C., Pilgreen, D., & Golebie, E. (2025). Measurement Instruments. In K. L. Pope, R. Arlinghaus, L. M. Hunt, A. J. Lynch, & B. T. van Poorten (Eds.), *Fish & Fisheries Series* (pp. 355-385). Springer. https://doi.org/10.1007/978-3-031-99739-6_12

Li, Z., Yu, J., Li, H., He, Y., Zhang-Li, D., Tu, S. et al. (2026). Beyond Self-Report: Bridging the Intention-Behavior Gap in Critical Thinking Assessment via Interpretable Multi-Agent System. In M. Liakata, V. P. Moreira, J. Zhang, & D. Jurgens (Eds.), *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (pp. 26849-26871). Association for Computational Linguistics. https://doi.org/10.18653/v1/2026.acl-long.1236

Mardia, K. V., Kent, J. T., & Taylor, C. C. (2024). *Multivariate Analysis*. John Wiley & Sons.

McAdams, D. P., & Dunlop, W. L. (2022). *The Person: A New Introduction to Personality Psychology*. John Wiley & Sons.

McNeish, D. (2018). Thanks Coefficient Alpha, We'll Take It from Here. *Psychological Methods, 23,* 412-433. https://doi.org/10.1037/met0000144

Minkin, V. A., Akimov, V. A., & Shelkanova, E. S. (2024). Patterns of Psychophysiological Responses to Multifactor Stimuli Associated with Multiple Intelligences and Personality Vices. In *Proceedings of the 7th International Open Science Conference: Modern Psychophysiology* (pp. 279-300).

Ntumi, S., Bulala, T., Yeboah, A., & Agbovor, D. (2026). Measuring the Immeasurable: Designing and Validating Assessments of Spiritual Intelligence as a Core Component of SEL in Faith-Based School Contexts in Ghana and Botswana. *Psychology in the Schools, 63,* 244-268. https://doi.org/10.1002/pits.70090

Piombo, M. A., La Grutta, S., Epifanio, M. S., Trombini, E., & Andrei, F. (2026). Differences between Children's Self-Reports and Teacher Ratings of Emotional Intelligence and Anxiety: A Cross-Informant Study. *Acta Psychologica, 264,* Article 106483. https://doi.org/10.1016/j.actpsy.2026.106483

Revelle, W., & Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the GLB: Comments on Sijtsma. *Psychometrika, 74,* 145-154. https://doi.org/10.1007/s11336-008-9102-z

Scheffer, D., & Heckhausen, H. (2025). Trait Theories of Motivation. In J. Heckhausen, & H. Heckhausen (Eds.), *Motivation and Action* (pp. 67-113). Springer. https://doi.org/10.1007/978-3-031-87947-0_3

Serapio-García, G., Safdari, M., Crepy, C., Sun, L., Fitz, S., Romero, P., Abdulhai, M., Faust, A., & Matarić, M. (2023). *Personality Traits in Large Language Models*. arXiv:2307.00184.

Sürücü, L., Yıkılmaz, İ., & Maşlakçı, A. (2024). Exploratory Factor Analysis (EFA) in Quantitative Researches and Practical Considerations. *Gümüşhane Üniversitesi Sağlık Bilimleri Dergisi, 13,* 947-965. https://doi.org/10.37989/gumussagbil.1183271

Tirri, K. (2023). Spirituality and Giftedness. *Gifted Education International, 39,* 73-79. https://doi.org/10.1177/02614294221129394

Tirri, K., & Nokelainen, P. (2008). Identification of Multiple Intelligences with the Multiple Intelligence Profiling Questionnaire III. *Psychology Science Quarterly, 50,* 206-221. https://www.psychologie-aktuell.com/fileadmin/download/PschologyScience/2-2008/08_Tirri.pdf

Tirri, K., & Nokelainen, P. (2011). Multiple Intelligences Profiling Questionnaire. In K.

Tirri, & P. Nokelainen (Eds.), *Measuring Multiple Intelligences and Moral Sensitivities in Education* (pp. 1-13). SensePublishers. https://doi.org/10.1007/978-94-6091-758-5_1

Tirri, K., Nokelainen, P., & Komulainen, E. (2013). Multiple Intelligences: Can They Be Measured? *Psychological Test and Assessment Modeling, 55,* 438-461.

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond G: Putting Multiple Intelligences Theory to the Test. *Intelligence, 34,* 487-502. https://doi.org/10.1016/j.intell.2006.02.004

Waterhouse, L. (2006). Multiple Intelligences, the Mozart Effect, and Emotional Intelligence: A Critical Review. *Educational Psychologist, 41,* 207-225. https://doi.org/10.1207/s15326985ep4104_1

Zumitzavan, V. (2026). Levels of Validity and Reliability. In V. Zumitzavan (Ed.), *Social Science Methodologies for Management Research*: From Philosophy to Practice in the AI Era (pp. 41-64). Springer. https://doi.org/10.1007/978-981-95-4318-2_4