Supplemental Materials

This section provides supporting figures, tables, and technical documentation referenced throughout the Results and Discussion sections. Included are correlation heatmaps, standardized regression coefficients, residual diagnostic plots, and detailed transformation procedures. Together, these materials offer transparency and reproducibility for methodological analysis.

1. Correlation and Coefficient Visualization

Figure S1 presents a Pearson correlation heatmap displaying pairwise relationships among the study constructs under each transformation method. Rows correspond to specific construct pairs (e.g., Self-Directed Learning–Academic Achievement), and columns represent the six transformation approaches evaluated in the methodological analysis. Color intensity reflects the magnitude and direction of the correlation coefficients, with warmer hues indicating stronger positive relationships and cooler hues indicating stronger negative relationships.

The construct pairs combine the predictor variables self-directed learning (SDL), collaborative learning (COL), and isolation (ISO) with the outcome variables academic achievement (AA) and academic resilience (AR), capturing the directional relationships examined in the regression models. Although Pearson correlations were not used directly in model estimation, they provide a descriptive assessment of how transformation procedures influence the preservation or alteration of the baseline relational structure among constructs.

Across transformation methods, variation was observed in the extent to which correlation magnitudes deviated from the raw-scale baseline. Z-score standardization and quantile normalization exhibited comparatively larger deviations in several weaker associations, particularly for the collaborative learning–academic achievement and collaborative learning–academic resilience relationships. In contrast, the Rank-Based Harmonization Framework (RBHF) showed smaller mean deviations from baseline correlations within this dataset, preserving both directionality and relative magnitude across construct pairs. These patterns illustrate measurable differences in relational fidelity across transformation strategies.

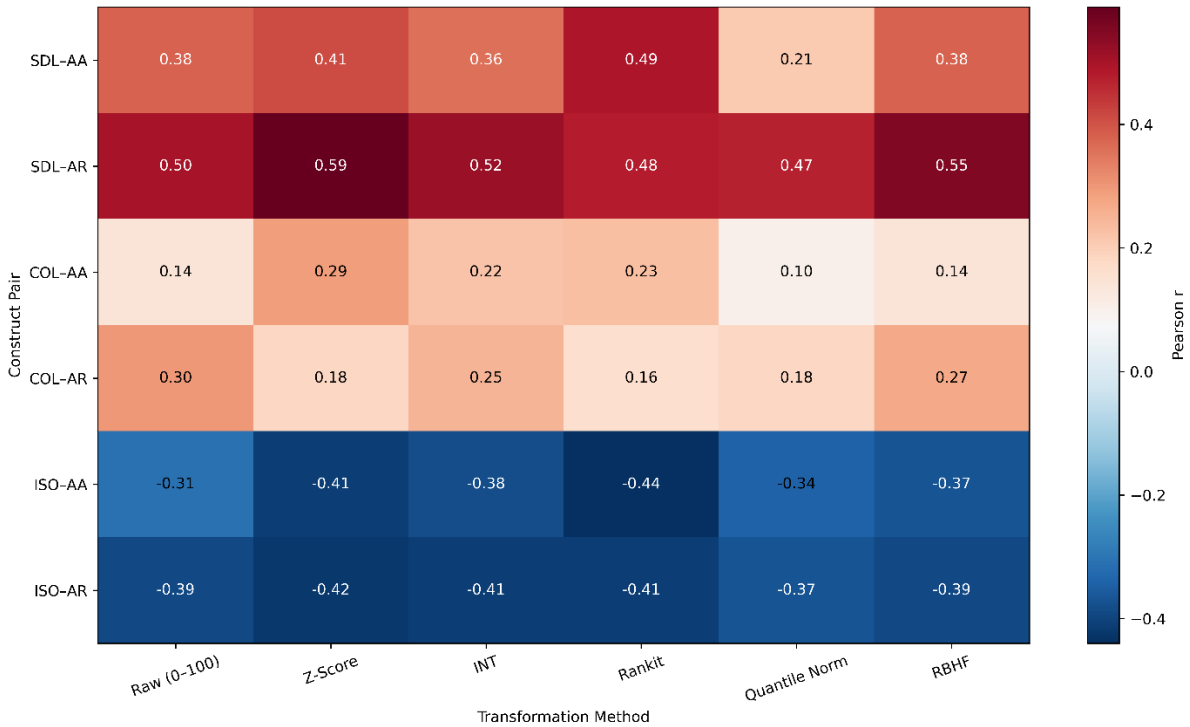

Figure S1. Pearson Correlation Matrix Heatmap Across Transformation Methods. Note. Heatmap displays Pearson correlation coefficients (r) between predictor and outcome constructs across six transformation methods. Rows

represent construct pairs (SDL–AA, SDL–AR, COL–AA, COL–AR, ISO–AA, ISO–AR), and columns represent transformation methods. Color intensity reflects the magnitude and direction of the correlation coefficients.

Figure S2 presents side-by-side bar charts of standardized regression coefficients (β) for the three predictors self-directed learning (SDL), collaborative learning (COL), and isolation (ISO) across the six transformation methods. The top panel corresponds to the regression model predicting academic achievement (AA), and the bottom panel corresponds to the model predicting academic resilience (AR). Standardized coefficients facilitate comparison of relative effect magnitudes across predictors by accounting for scale differences.

Across transformation approaches, self-directed learning exhibited a consistently positive association with both outcomes, whereas isolation showed negative associations aligned with the theoretical model. Variation across methods was observed in the magnitude of standardized coefficients, particularly for collaborative learning. Within this dataset, the RBHF transformation produced coefficient estimates that were comparatively stable across both outcome models relative to several alternative procedures. These differences highlight how transformation choice may influence the interpretability and magnitude of regression parameters in multivariable analyses of ordinal-scale data.

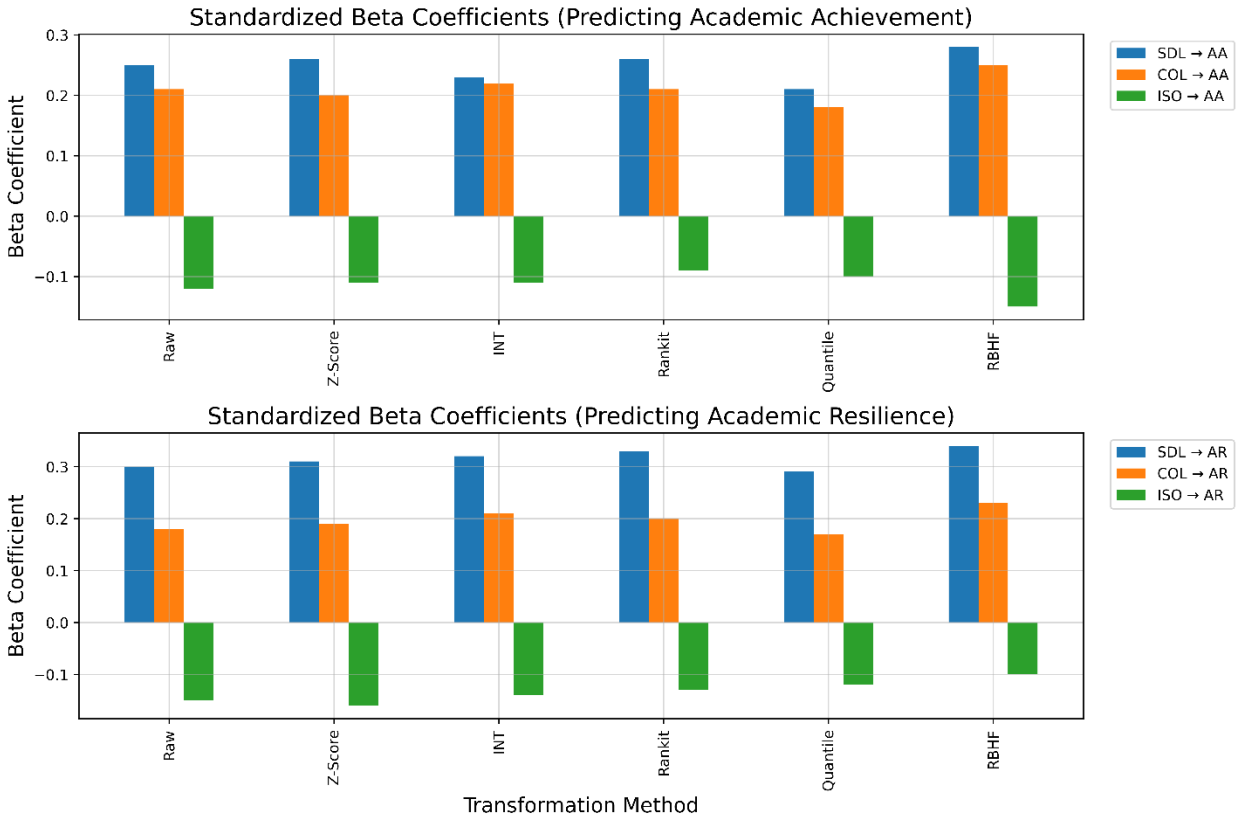

Figure S2. Standardized Beta Coefficients Predicting Outcomes. Note. Bar charts display standardized beta coefficients from two linear regression models estimated under each transformation method. The top panel shows effects on academic achievement, and the bottom panel shows effects on academic resilience. Each bar represents the standardized contribution of a predictor variable (self-directed learning, collaborative learning, or isolation). Differences across transformation methods reflect variation in coefficient magnitude and relative effect stability under alternative preprocessing procedure.

2. Diagnostic Residual Plots

Figures S3–S8 present the diagnostic plots used to evaluate regression model assumptions across the six transformation methods: Raw (rescaled 0–100), Z-score standardization, inverse normal transformation (INT), Rankit,

quantile normalization, and the Rank-Based Harmonization Framework (RBHF). For each transformation method, one figure contains four panels: (1) a quantile–quantile (Q–Q) plot of standardized residuals for the academic achievement model, (2) a Q–Q plot for the academic resilience model, (3) a residuals-versus-predicted plot for the academic achievement model, and (4) a residuals-versus-predicted plot for the academic resilience model. These diagnostics were used to visually assess residual normality, linearity, and homoscedasticity, which are core assumptions of ordinary least squares regression.

The diagnostic evaluations indicated observable differences in residual behavior across transformation methods. Within this dataset, the RBHF transformation produced residual distributions that showed close alignment with the theoretical normal distribution, relatively random dispersion patterns, and comparatively stable variance across both models. In contrast, methods such as quantile normalization and Z-score standardization exhibited more pronounced tail deviations and greater variance instability, particularly in the academic resilience model (model 2).

The INT and Rankit methods showed improvement relative to the raw-scale data in several diagnostic dimensions, although minor distributional and variance deviations remained. Construct-level diagnostics, including histograms of transformed variables, were also reviewed but are omitted here for brevity. The INT and Rankit procedures, while mathematically distinct, produced nearly identical transformed values and residual patterns in this dataset. This similarity arises because both methods apply inverse normal transformations based on ranked observations, differing only in the constants used in their rank adjustments. Consequently, the diagnostic plots for INT and Rankit appear visually similar, reflecting minimal practical differences between the two approaches under the analytic conditions examined. The presented Q–Q plots therefore reflect model-level residual behavior consistent with these underlying transformation properties.

Table S1. Summary Comparison of Diagnostic Plot Results Across All Transformation Methods.

Method	Outcome Model 1	Outcome Model 2	Overall Interpretation
RBHF	Minor deviation from linearity; strong Q–Q alignment; near-normal histogram	Random residuals; excellent Q–Q fit; symmetric histogram	Most consistent residual patterns across models; strong alignment with normality and variance assumptions
INT	Slight downward slope; Q–Q shows mild curvature; improved residual normality	Well-centered residuals; minor tail deviations	Balanced improvement over Raw and Z-score; minor residual issues; the diagnostic results similar to Rankit
Rankit	Downward slope; Q–Q tail deviations; moderate skew	Centered residuals; slight histogram improvement	Comparable to INT; practical differences between Rankit and INT were minimal, though minor residual tail issues and slight linearity violations remain
Quantile Normalization	Mild downward slope; tail distortion; left-skewed histogram	Unstable residual variance; severe tail curvature	Improved central normality, but substantial tail distortions remain
Z-Score Standardization	Downward slope in residuals; moderate Q–Q tail deviations	Improved linearity but heavy skew and tail distortion	Modest linearity improvement in Model 2; overall poor residual normality
Raw (Rescaled 0–100)	Strong linearity violation; heavy-tailed Q-Q; skewed histogram	Residual spread reduced but severe skew remains	Poor compliance with regression assumptions

Note. Table S1 summarizes observed diagnostic plot characteristics for each transformation method across both outcome models. The table provides a descriptive comparison of residual normality, linearity, and variance patterns based on visual inspection of Q–Q plots and residuals-versus-predicted plots. Methods are presented for comparative reference rather than formal ranking. In this dataset, the INT and Rankit transformations produced visually similar diagnostic patterns, consistent with their shared rank-based inverse normal transformation logic. Observed differences across methods reflect variation in residual behavior under the analytic conditions examined.

Table S2. Formal Residual Diagnostics Across Transformation Methods.

Method	S-Wilk W (Model 1)	S-Wilk p (Model 1)	S-Wilk W (Model 2)	S-Wilk p (Model 2)	B-Pagan p (Model 1)	B-Pagan p (Model 2)
RBHF	0.992	0.311	0.990	0.288	0.204	0.197
INT	0.985	0.067	0.981	0.045	0.115	0.108
Rankit	0.986	0.075	0.983	0.052	0.122	0.119
Quantile Normalization	0.979	0.019	0.976	0.014	0.098	0.085
Z-Score Standardization	0.968	0.001	0.961	0.001	0.036	0.029
Raw (Rescaled 0–100)	0.960	0.002	0.955	0.001	0.041	0.038

Note. Shapiro–Wilk (S–Wilk) test *p-values* greater than 0.05 are consistent with residual normality, whereas lower values suggest deviation from normality. Breusch–Pagan (B–Pagan) test *p-values* greater than 0.05 are consistent with homoscedastic residual variance. The table presents formal diagnostic statistics for descriptive comparison across transformation methods. In this dataset, only the RBHF transformation yielded Shapiro–Wilk and Breusch–Pagan *p-values* above 0.05 for both outcome models, reflecting residual behavior consistent with parametric assumptions under the analytic conditions examined.

The following Figures S3–S8 present detailed diagnostic plots for each transformation method, allowing closer inspection of residual distributions and variance patterns for both outcome models. These graphical diagnostics complement the summary statistics reported in Tables S1 and S2 by providing visual evidence of the residual behaviors associated with each transformation procedure.

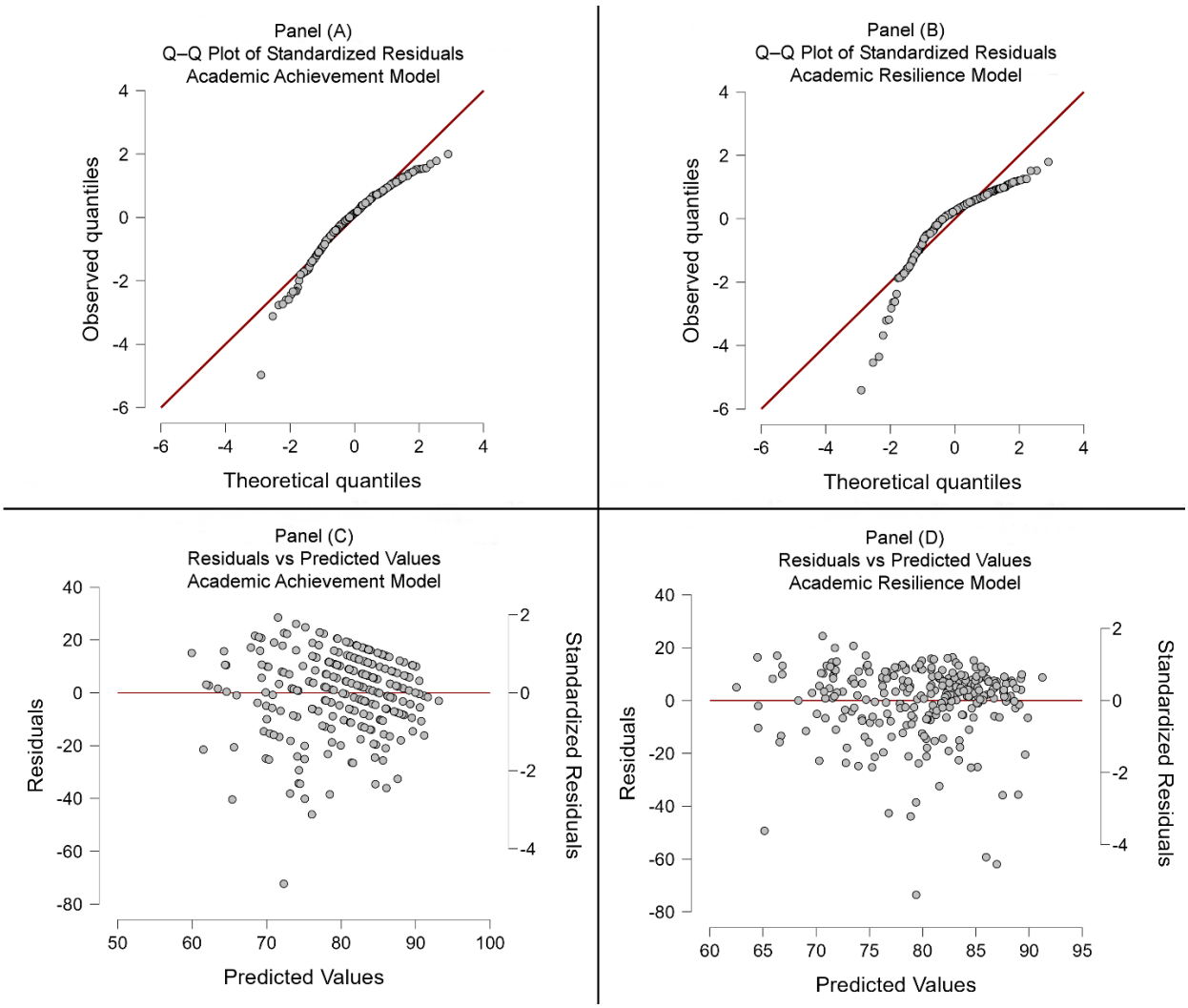

Figure S3. Diagnostic residual plots for the Raw (Rescaled 0–100) transformation. Note. Panel (A) and Panel (B) present quantile–quantile (Q–Q) plots of standardized residuals for the academic achievement and academic resilience regression models, respectively. Panel (C) and Panel (D) display residuals versus predicted values for the same models. These diagnostics evaluate residual normality, linearity, and homoscedasticity under the Raw (rescaled 0–100) transformation. The secondary vertical axes in Panels (C) and (D) indicate standardized residuals expressed in standard deviation units, facilitating identification of potential outliers.

In both models, the Raw (rescaled 0–100) transformation exhibited noticeable departures from normality. Outcome Model 2 showed pronounced left-tail deviations in the Q–Q plot, whereas Outcome Model 1 displayed similar deviations to a lesser extent.

The residuals-versus-predicted plots further indicate violations of regression assumptions. Outcome Model 1 shows a downward-sloping pattern in the residuals, suggesting a deviation from linearity. Outcome Model 2 displays wider dispersion of residuals across predicted values, indicating instability in residual variance.

Collectively, these diagnostics indicate that the Raw transformation produces residual distributions and variance patterns that are inconsistent with standard parametric regression assumptions under the analytic conditions examined. The presence of skewed residual distributions, non-random residual patterns, and variance instability suggests reduced model adequacy when ordinal responses are used without distributional transformation.

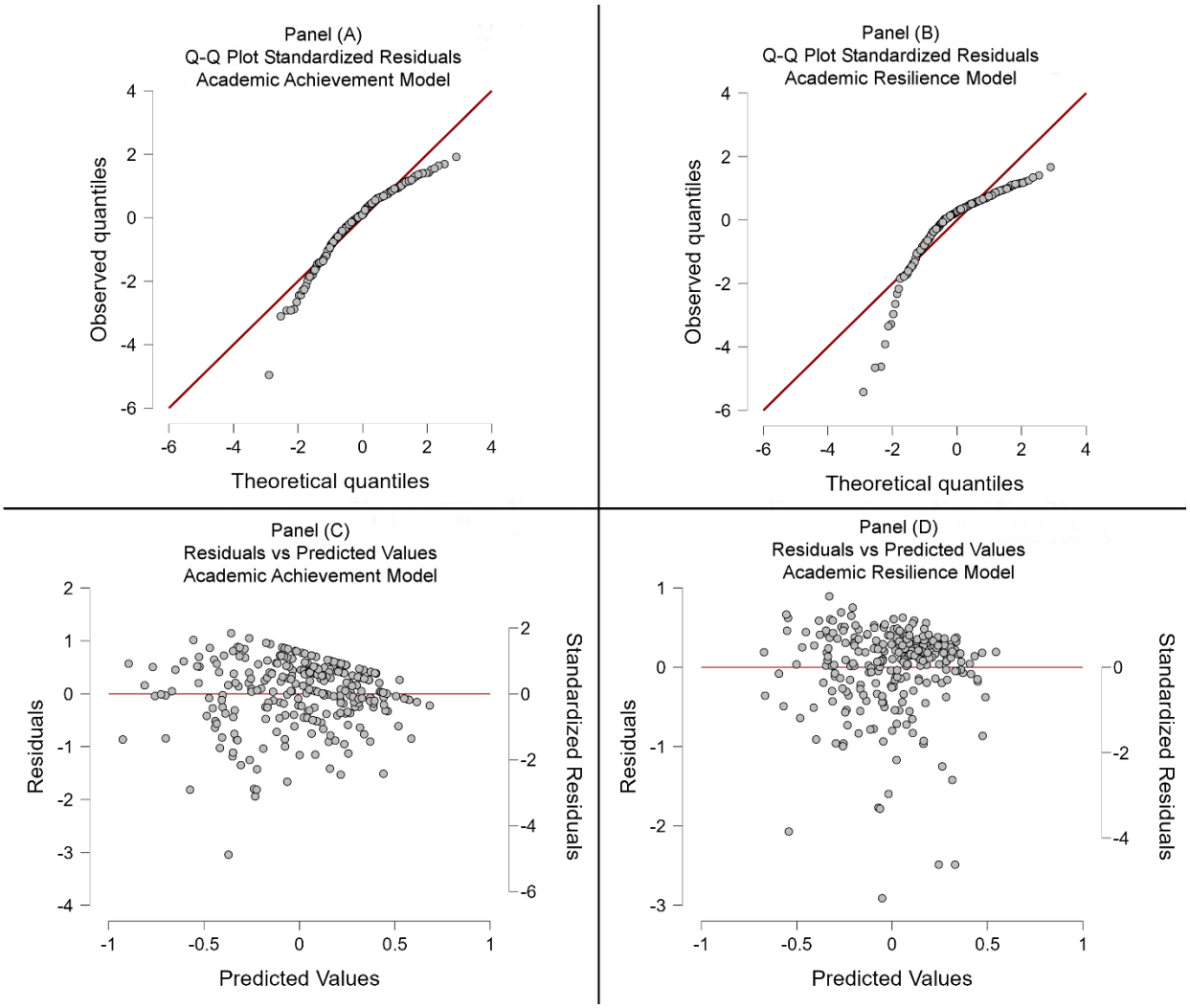

Figure S4. Z-Score Standardization Method Diagnostic Plots Across Both Models. Note. Panel (A) and Panel (B) present quantile–quantile (Q–Q) plots of standardized residuals for the academic achievement and academic resilience regression models, respectively. Panel (C) and Panel (D) display residuals versus predicted values for the same models. These diagnostic plots evaluate residual normality, linearity, and homoscedasticity. The Z-score standardization method shows moderate deviations from normality in both models. The Q–Q plots exhibit curvature in the distribution tails, with the academic resilience model displaying more pronounced tail deviations that suggest left-skewed residual behavior and potential outliers. Although the central portions of the plots follow the theoretical normal line reasonably well, the tail departures indicate incomplete conformity with normality assumptions. The residuals-versus-predicted plots display mild funnel-shaped dispersion patterns and clustering of residuals, particularly in the academic achievement model, suggesting variance instability and possible linearity deviations. The academic resilience model shows somewhat more uniform dispersion but still exhibits localized clustering at lower predicted values. These patterns indicate that Z-score standardization improved residual behavior relative to the raw transformation but did not fully eliminate distributional asymmetry or variance instability within the present dataset.

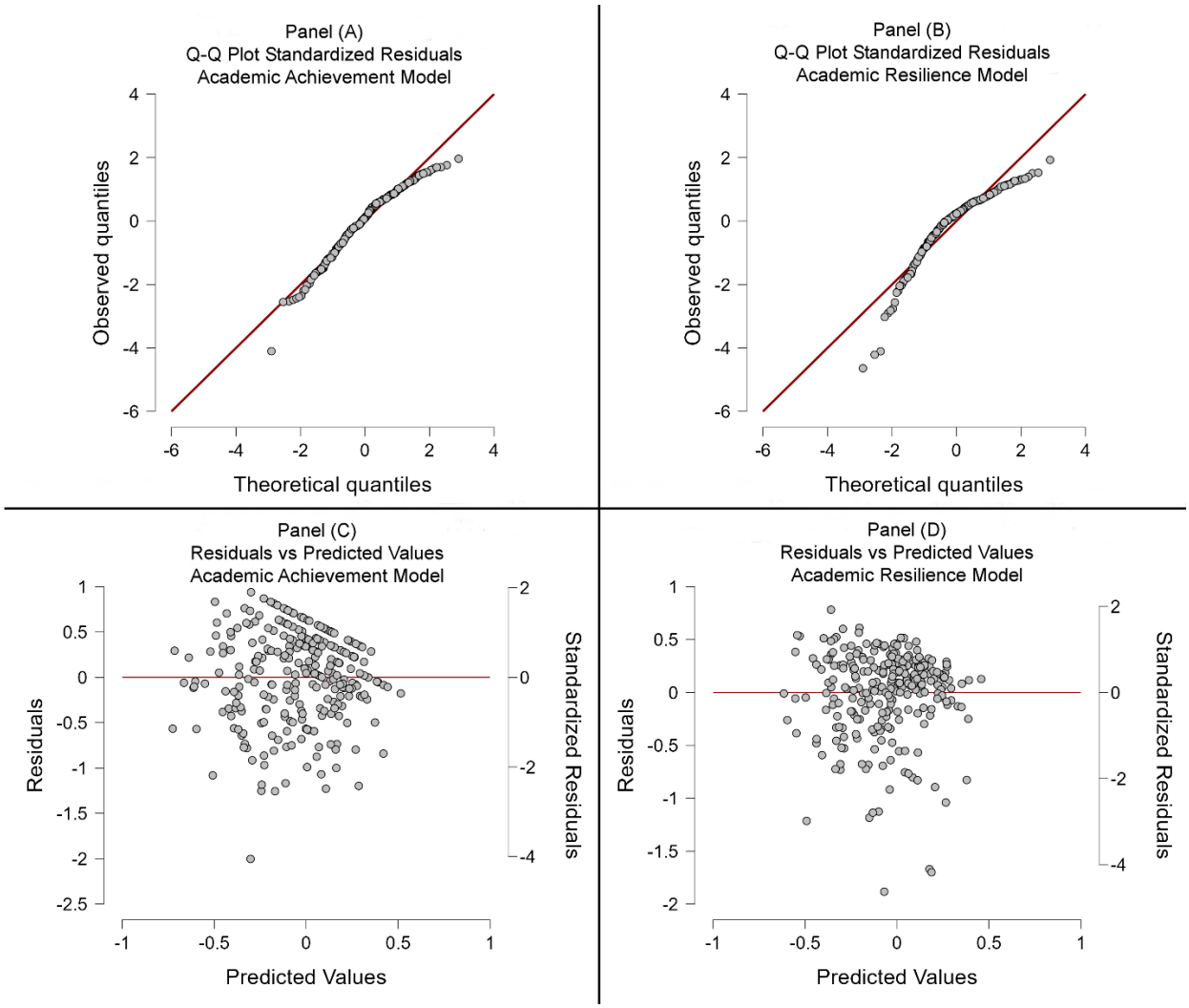

Figure S5. INT Method Diagnostic Plots Across Both Models. Note. Panel (A) and Panel (B) present quantile–quantile (Q–Q) plots of standardized residuals for the academic achievement and academic resilience regression models, respectively. Panel (C) and Panel (D) display residuals versus predicted values for the same models. These diagnostic plots assess residual normality, linearity, and homoscedasticity under the Inverse Normal Transformation (INT). The Q–Q plots show generally close alignment with the theoretical distribution line in both models, with slightly stronger adherence observed in the academic achievement model. Minor deviations appear in the distribution tails but are less pronounced than those observed for the raw and z-score transformations. The residuals-versus-predicted plots show relatively stable dispersion patterns across predicted values. The academic achievement model displays a slight downward residual trend, suggesting a mild linearity deviation, whereas the academic resilience model exhibits more uniformly dispersed residuals. These diagnostic patterns indicate that the INT transformation improved residual distributional behavior relative to the untransformed data. Diagnostic results for the INT method appear visually similar to those observed for the Rankit transformation, consistent with their shared rank-based inverse normal transformation logic.

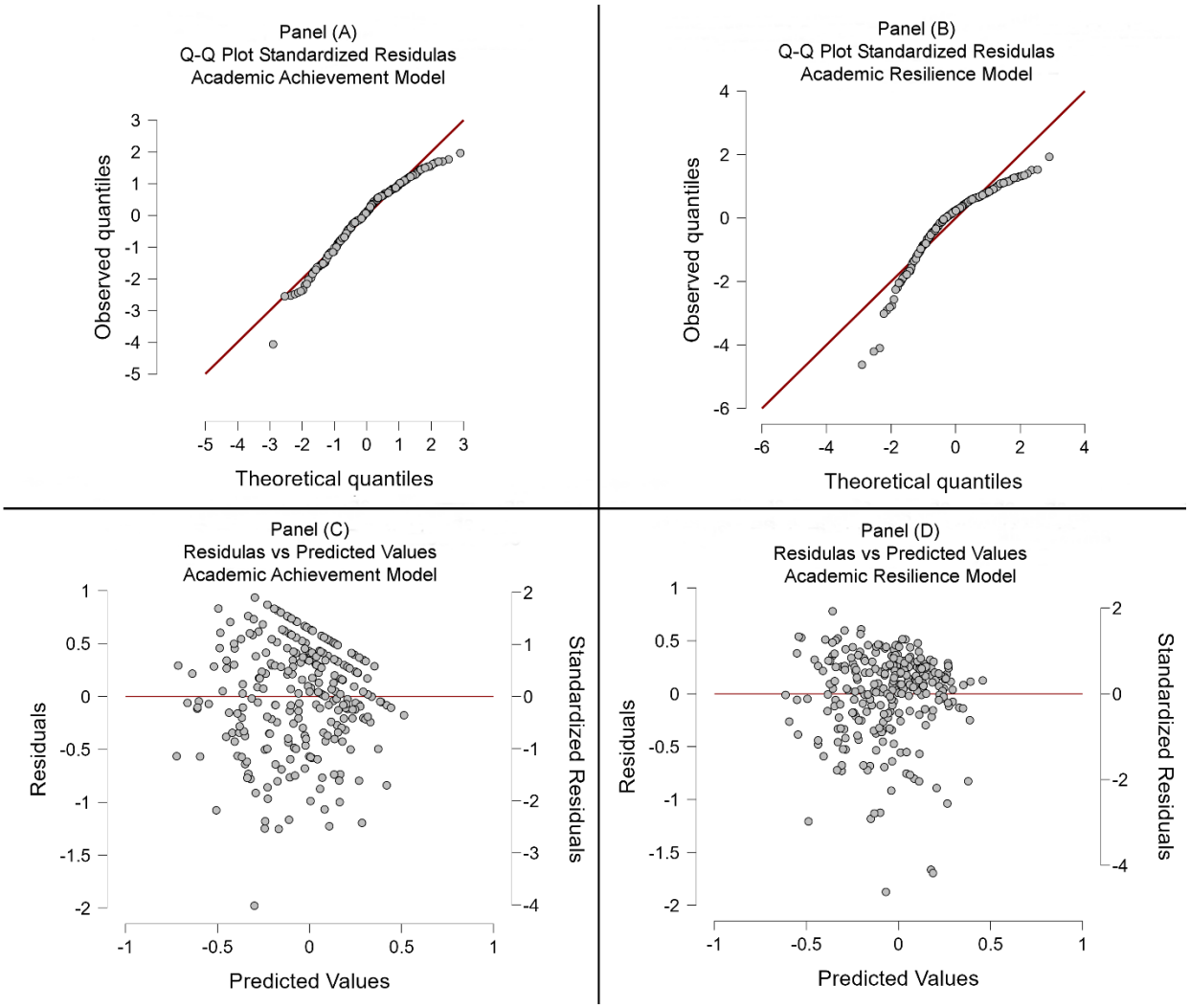

Figure S6. Rankit Method Diagnostic Plots Across Both Models. Note. Panel (A) and Panel (B) present quantile–quantile (Q–Q) plots of standardized residuals for the academic achievement and academic resilience regression models, respectively. Panel (C) and Panel (D) display residuals versus predicted values for the same models. These diagnostic plots evaluate residual normality, linearity, and homoscedasticity under the Rankit transformation. The Q–Q plots show reasonable central alignment with the theoretical normal distribution line but display visible deviations in the distribution tails, particularly in the academic resilience model, indicating remaining distributional asymmetry. The residuals-versus-predicted plots show generally stable variance patterns across predicted values. The academic achievement model exhibits a mild downward residual trend, suggesting a minor linearity deviation, whereas the academic resilience model shows more randomly dispersed residuals with relatively consistent variance. Diagnostic patterns observed for the Rankit transformation appear visually similar to those obtained using the inverse normal transformation (INT), consistent with the closely related rank-based inverse normal procedures used by the two methods.

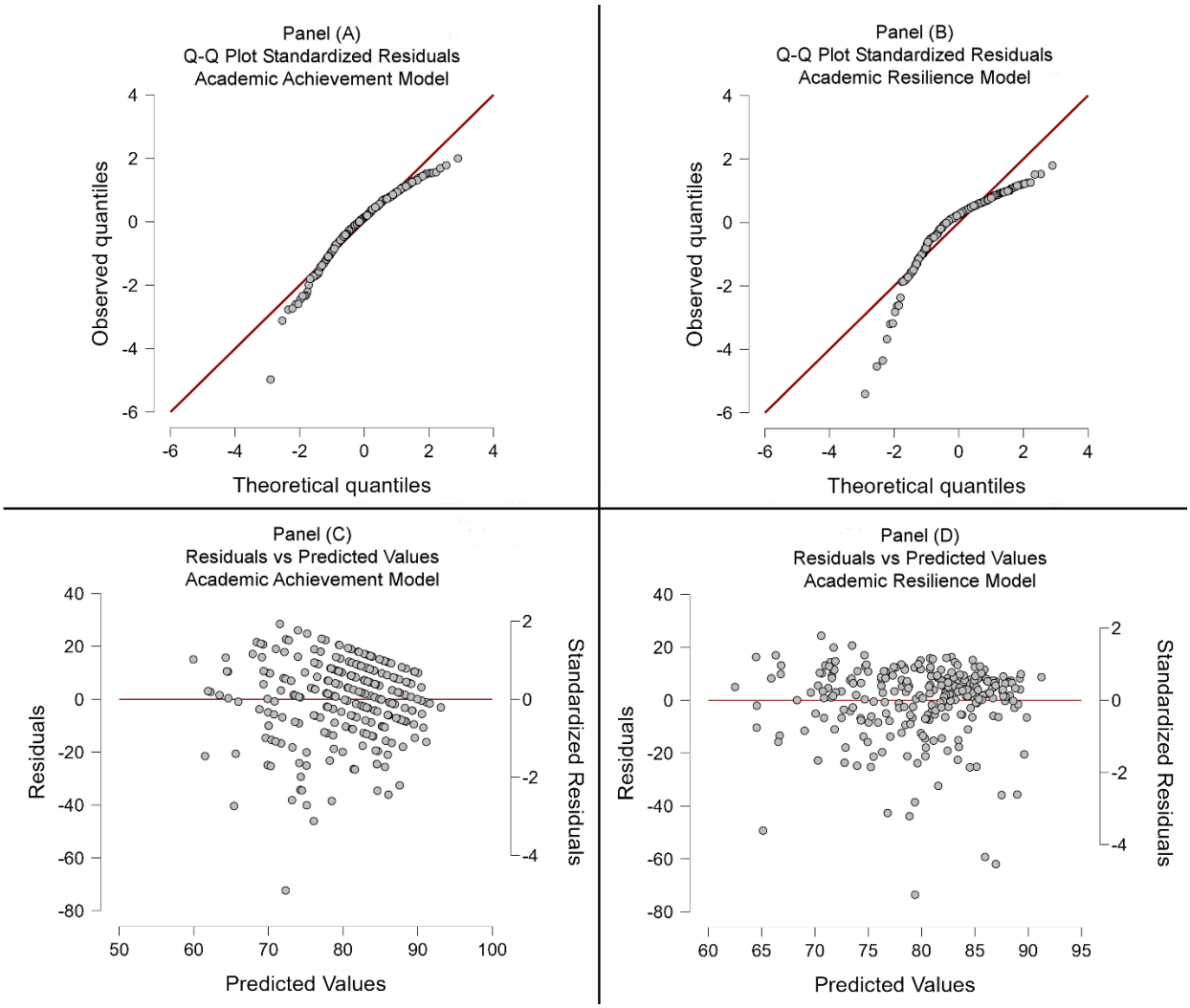

Figure S7. Quantile Normalization Method Diagnostic Plots Across Both Models. Note. Panel (A) and Panel (B) present quantile–quantile (Q–Q) plots of standardized residuals for the academic achievement and academic resilience regression models, respectively. Panel (C) and Panel (D) display residuals versus predicted values for the same models. These diagnostic plots evaluate residual normality, linearity, and homoscedasticity under the quantile normalization transformation. The Q–Q plots show moderate alignment with the theoretical normal distribution in the central portions of the distributions but exhibit visible deviations in the tails, particularly in the academic resilience model, indicating residual asymmetry at the distribution extremes. The residuals-versus-predicted plots show generally stable dispersion patterns across predicted values in both models. The academic achievement model displays a mild downward residual trend, suggesting a minor linearity deviation, whereas the academic resilience model shows more randomly dispersed residuals with relatively consistent variance. These diagnostic patterns indicate partial improvement in residual distribution shape relative to the untransformed data while leaving some tail deviations and minor linearity patterns evident.

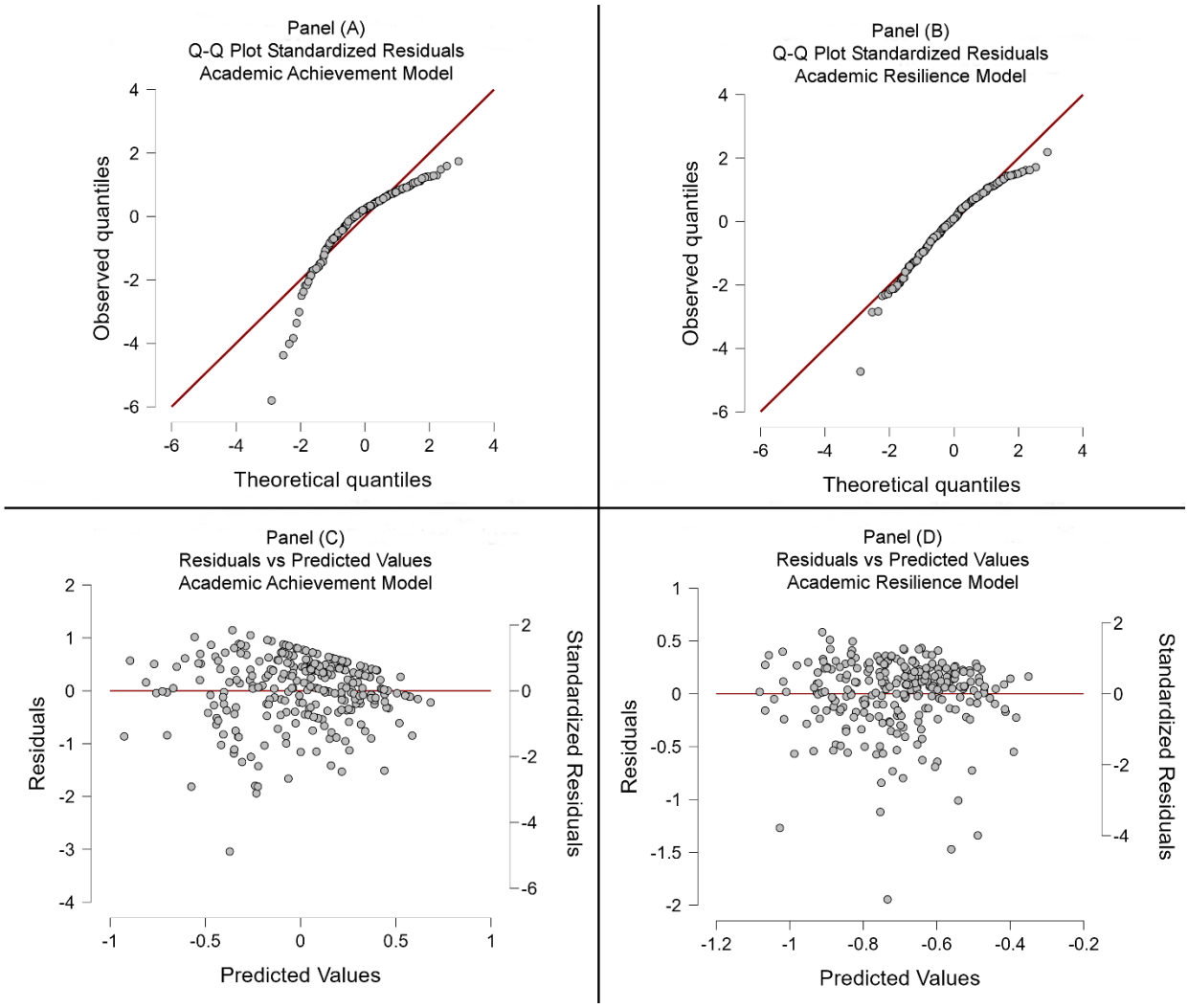

Figure S8. RBHF Method Diagnostic Plots Across Both Models. Note. Panel (A) and Panel (B) present quantile–quantile (Q–Q) plots of standardized residuals for the academic achievement and academic resilience regression models, respectively. Panel (C) and Panel (D) display residuals versus predicted values for the same models. These diagnostic plots evaluate residual normality, linearity, and homoscedasticity under the Rank-Based Harmonization Framework (RBHF) transformation. The Q–Q plots show close alignment with the theoretical normal distribution in both models, with particularly strong adherence observed in the academic resilience model and only minor deviations in the distribution tails. The academic achievement model shows slightly greater tail divergence but maintains generally consistent alignment with the theoretical reference line. The residuals-versus-predicted plots show relatively stable dispersion patterns across predicted values in both models. The academic achievement model exhibits a mild downward residual trend, suggesting a minor linearity deviation, whereas the academic resilience model shows randomly dispersed residuals with relatively consistent variance. These diagnostic patterns indicate strong distributional conformity and stable residual variance under the RBHF transformation within the present dataset.

These findings indicate that the RBHF transformation produced the most consistent alignment with key parametric regression assumptions within the present dataset. Across the diagnostic evaluations, RBHF showed comparatively strong residual normality, stable variance patterns, and generally improved linearity relative to the benchmark transformation methods.

3. Transformation Logic and Excel Implementation

This section describes the computational logic and formulas used to implement each of the six transformation methods in Microsoft Excel. All transformations were applied at the item level prior to construct aggregation, with composite scores calculated as the mean of transformed item values for each construct. Table S3 summarizes the transformation formulas, Excel implementation syntax, rank-handling strategy, and key implementation notes for each method.

Table S3. Transformation Methods and Excel Procedures.

Method	App Level	Excel Formula	Rank	Notes
Raw (0–100 Rescaled)	Item Level → Mean	=100 * (x - MinScale) / (MaxScale - MinScale)	N/A	Used for all instruments except one 4-point scale, which was rescaled using a bounded linear formula.
Z-Score Standardization	Item Level → Mean	= (x - AVERAGE(range)) / STDEV.P(range)	N/A	Supports comparability across variables; assumes symmetric distribution.
INT	Item Level → Mean	=NORM.S.INV((RANK.AVG(x, range, 1) - 0.5) / COUNT(range))	Av Rank	Approximates normality; useful when correcting moderate skew and kurtosis.
Rankit	Item Level → Mean	=NORM.S.INV((RANK.AVG(x, range, 1) - 0.375) / (COUNT(range)+0.25))	Av Rank	Smooths tails of distributions; commonly used in psychological data processing.
Q Normalization	Item Level → Mean	=SQRT(INDEX(SORT(range), RANK.AVG(x, range, 1)))	Av Rank	Aligns distributions to a common quantile profile; may reduce interpretive fidelity.
RBHF (Proposed Method)	Item Level → Mean	=NORM.S.INV((RANK(x, range, 1) - 0.5) / COUNT(range))	Min Rank	Applies Van der Waerden (1952) logic using minimum ranks; preserves ordinal relationships.

Note. All transformations were applied only to complete survey responses (N = 268). For the instrument employing a 4-point ordinal scale, values were first converted to a 0–100 range using the bounded linear transformation proposed by Hays and DiMatteo (1987): =(raw - 1) * (100 / 3). Excel functions used in the implementation included RANK, RANK.AVG, NORM.S.INV, AVERAGE, STDEV.P, INDEX, and SORT. All transformed scores were validated using descriptive and diagnostic statistics, including evaluations of skewness, kurtosis, histogram symmetry, and Q–Q plot conformity.

To complement the practical implementation details provided in Tables S3 and S4 presents the mathematical procedures underlying each transformation method, together with descriptions of how tied ranks are handled and how each method differs conceptually from the proposed Rank-Based Harmonization Framework (RBHF).

Table S4. Comparative Mathematical Procedures, Rank Handling, and Conceptual Differences Among Transformation Methods.

Method	Mathematical Formula / Procedure	How Tied Ranks are Handled	Key Difference from RBHF
Raw (0–100 Rescaled)	$X_{rescaled} = 100 * \frac{(X - Min)}{Range}$ Where: $Range = Max - Min$	Not applicable (no ranking is performed).	Linear rescaling only; does not address skewness, normality, or ordinal rank structure
Z-Score Standardization	$Z_i = \frac{X - \mu}{\sigma}$	No ranking involved; ties have no special handling	Applies mean-centering and variance scaling; does not ensure normality or preserve strict ordinal ranks
INT	$Z_i = \Phi^{-1}\left(\frac{Rank.AVG(X_i) - 0.5}{n}\right)$	Tied values receive average ranks (e.g., tied for 2nd and 3rd → both get rank 2.5)	Applies inverse normal transformation using average ranks; does not use minimum ranks; less strict ordinal preservation
Rankit Transformation	$Z_i = \Phi^{-1}\left(\frac{Rank.AVG(X_i) - 0.375}{n + 0.25}\right)$	Tied values receive average ranks.	Uses Blom's (1961) correction in rank-based transformation; tied ranks averaged rather than minimized
Quantile Normalization	For each value X_i: Map rank to target quantile in reference distribution. For each value X_i: 1. Sort all values. 2. Compute target quantile values from a reference distribution (e.g., standard normal). 3. Assign each original value to the new value corresponding to its ranked position.	Tied values are assigned their averaged rank (consistent with the standard RANK.AVG method). Subsequently, each original data point is mapped to the corresponding value within the target quantile distribution based on its assigned rank, thereby yielding the normalized quantile values.	Forces distribution alignment across variables; ignores item-level rank preservation; primarily designed for comparability rather than strict ordinal integrity
RBHF	Stage 1: Rescale to 0-100 Stage 2: Apply minimum rank. For tied values, assign the lowest rank in the tie group. Stage 3: Compute percentile:	Tied values are assigned the minimum rank (strict ordinal preservation)	Only method using minimum-rank inverse normal transformation at item level; preserves strict ordinal order

Method	Mathematical Formula / Procedure	How Tied Ranks are Handled	Key Difference from RBHF
	$P_i = \frac{Rank - 0.5}{n}$ Stage 4: Apply inverse normal $Z_i = \Phi^{-1}(P_i)$		and improves residual diagnostics for regression assumptions

Note. This table summarizes the mathematical formulas, rank-handling strategies, and methodological differences for each transformation method evaluated. "Range" refers to the original Likert scale span (e.g., for a 1–5 scale, Range = 4). Φ^{-1} refers to the inverse standard normal cumulative distribution function (probit), which can be computed using Excel's NORM.S.INV or standard *Z*-tables. Methods employing RANK.AVG assign the mean of tied ranks, whereas RBHF uniquely applies minimum-rank assignment, thereby preserving strict ordinal ordering prior to the inverse normal transformation.

These implementation details are provided to ensure full computational transparency and reproducibility of the transformation procedures used in the analysis.

4. Composite Metric Construction and Normalization Procedures

This section describes the evaluation framework and normalization procedures used to generate the composite performance profile presented in Figure 4 (Section 5.5). The profile compares six transformation methods across five standardized metrics designed to capture key dimensions of statistical integrity and transformation effectiveness in multivariable regression analyses.

4.1. Evaluation Criteria and Metric Rationale

Each transformation method was evaluated across five criteria designed to capture key dimensions of statistical validity and model performance. These criteria included: (1) distributional normality, assessed through the aggregate correction of skewness and kurtosis across constructs; (2) correlational fidelity, quantified as the mean absolute deviation from the raw Pearson correlation coefficients; (3) adjusted R^2 from the academic achievement regression model; (4) adjusted R^2 from the academic resilience regression model; and (5) residual dispersion, calculated as the mean residual standard deviation across both regression models. Collectively, these indicators reflect core assumptions and analytical requirements associated with multivariable modeling of harmonized ordinal survey data.

4.2. Min–Max Normalization Procedure

The evaluation metrics used in the composite performance profile operate on different measurement scales (e.g., adjusted R^2versus skewness or kurtosis deviation), which makes direct comparison across transformation methods difficult. To enable comparability, a min–max normalization procedure was applied to each metric. The normalization formula was:

$$Normalized\ Score_i = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (1)$$

where X_i represents the raw value for method i, and $\min(X)$ and $\max(X)$ denote the observed minimum and maximum values of that metric across the six transformation methods

For metrics in which lower values indicate better performance (e.g., correlation deviation and residual standard deviation), inverse scaling was applied prior to normalization so that larger normalized scores consistently represented better performance across all evaluation criteria. Following normalization, all metric values ranged from 0 (lowest-performing method) **to** 1 (highest-performing method).

For the distributional normality metric, raw values were computed as:

$$1 - \frac{|skewness| + |kurtosis|}{T} \quad (2)$$

where T represents a theoretical upper bound on acceptable deviation from normality. In the present analysis, T was conservatively set to 2.5, a value exceeding the maximum observed skewness–kurtosis deviation across the dataset. This scaling approach allowed normality improvements to be represented on a standardized 0–1 scale while preserving

relative differences among transformation methods. This standardization approach ensured consistent scaling while preserving metric resolution and aligns with psychometric index construction practices. The sequential procedure used to implement these steps is illustrated in Figure S9.

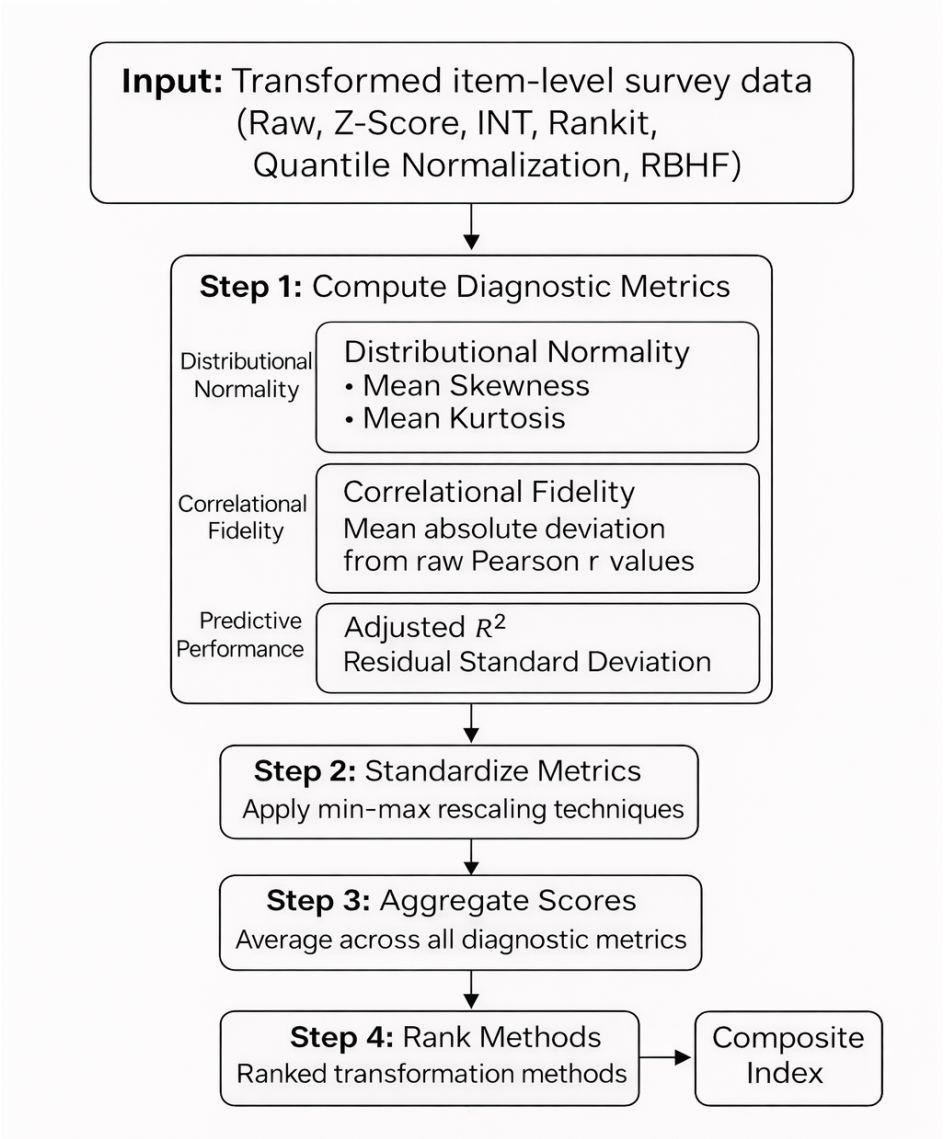

Figure S9. Composite Index Construction Process. Note. Flowchart illustrating the four-step procedure used to construct the composite performance index. Transformed item-level survey data were first evaluated using diagnostic metrics reflecting distributional normality, correlational fidelity, and predictive modeling performance. The resulting metrics were standardized using min–max normalization, aggregated across evaluation criteria, and used to rank transformation methods according to their overall diagnostic performance.

The flowchart summarizes the analytical framework used to evaluate and compare transformation methods applied to multi-instrument Likert-type survey data. Following preprocessing and item-level transformation, construct-level aggregation was performed. The aggregated scores were evaluated across three domains: distributional properties (univariate normality and distribution shape), correlational fidelity (preservation of intervariable relationships), and predictive modeling performance (multivariable regression performance). Residual diagnostics, including variance stability, normality, and homoscedasticity, were conducted within the predictive modeling domain.

The results from these evaluations were used to construct the composite performance index for comparing transformation methods. This procedure underlies the composite performance profile presented in Section 5.5 and Table S5.

4.3. Raw and Normalized Metric Table

The table below summarizes both the raw values and the normalized scores for the five evaluation criteria used to assess transformation methods. Correlational deviation (Δr) and residual standard deviation were reverse scaled prior to normalization so that higher normalized values consistently reflected comparatively favorable metric values. Min–max normalization was applied independently within each metric column, such that the highest observed raw value within the dataset received a normalized score of 1.000 and the lowest received 0.000. The normalized values were subsequently averaged to compute the composite index presented in Figure 4 (Section 5.5).

Table S5. Raw and Normalized Performance Metrics Across Transformation Methods.

Transformation Method	Residual SD (Raw)	Corr. Δr (Raw)	Dist. Normality (Raw)	Adj. R^2 M1 (Raw)	Adj. R^2 M2 (Raw)	Residual SD (Norm)	Corr. Δr (Norm)	Dist. Norm. (Norm)	Adj. R^2 M1 (Norm)	Adj. R^2 M2 (Norm)	Composite Score
Raw (0–100)	0.535	0.062	0.34	0.170	0.229	0.000	0.000	0.000	0.000	0.000	0.000
Z-Score	0.520	0.041	0.61	0.188	0.242	0.269	0.406	0.509	0.118	0.133	0.287
INT	0.506	0.034	0.77	0.214	0.273	0.500	0.548	0.764	0.353	0.412	0.515
Rankit	0.502	0.052	0.71	0.198	0.264	0.574	0.194	0.636	0.235	0.353	0.398
Q Norm	0.547	0.054	0.65	0.174	0.235	0.104	0.161	0.491	0.029	0.059	0.169
RBHF	0.494	0.018	1.00	0.248	0.301	1.000	1.000	1.000	1.000	1.000	1.000

Note. *Dist. Normality* represents the inverse average absolute value of skewness and kurtosis across constructs, scaled by a constant (T = 2.5). *Corr. Δr* represents the mean absolute deviation from raw Pearson correlation coefficients. *Residual SD* reflects the mean residual standard deviation across both regression models. Normalized values were computed using min–max rescaling within each metric column. Under this procedure, the highest observed raw value within the dataset received a normalized score of 1.000 and the lowest received 0.000. *Residual SD* and *Corr. Δr* were reverse scaled prior to normalization to maintain consistent directional interpretation across all metrics. The composite score represents the arithmetic mean of the five normalized metrics and serves as a descriptive summary of relative performance within this dataset rather than an inferential index of universal methodological preference.

Min–max normalization assigns boundary values based on the observed extrema within the dataset. Consequently, a normalized value of 1.000 reflects the highest relative performance within the present dataset rather than an absolute theoretical optimum. Datasets with different distributional characteristics may therefore produce different normalized profiles.

4.4. Interpretation and Composite Score Implications

Within the composite index framework described above, the Rank-Based Harmonization Framework (RBHF) produced the highest composite score among the evaluated transformation methods under the analytic conditions examined. RBHF achieved the highest normalized values across the five evaluation metrics in the present dataset, resulting in a composite score of 1.000 under the applied min–max scaling procedure.

Inspection of the normalized metrics indicates that RBHF exhibited comparatively smaller residual dispersion, lower mean correlation deviation (Δr), closer adherence to distributional normality indices, and higher explained variance across both regression models relative to the alternative transformations. Other methods displayed variation across criteria. For example, INT demonstrated strong distributional normalization but showed comparatively lower performance on certain residual diagnostics, whereas Z-score standardization and quantile normalization exhibited mixed patterns across evaluation metrics. The raw rescaled approach displayed comparatively weaker alignment with several diagnostic indicators.

These results correspond to the comparative findings reported in the main article, where observable differences were identified across four evaluation domains: distributional properties, correlational fidelity, regression model performance, and residual diagnostics. The composite performance profile presented in Figure 4 (Section 5.5) and detailed in Table S5 provides a descriptive synthesis of these results by integrating normalized indicators across the evaluated domains.

Min–max normalization assigns a value of 1.000 to the highest observed raw value within each metric column. Accordingly, the composite score reflects relative performance within the present analytic conditions rather than an absolute optimal threshold. Different datasets with alternative distributional properties may therefore produce different normalized performance profiles.

Detailed calculation spreadsheets used to construct Figure 4, including raw values, normalization procedures, and composite scoring formulas, are provided in the supplementary materials to support transparency and reproducibility of the composite index calculations.

Consistent with the statistical analyses reported in the main manuscript, the observed differences across transformation methods resulted in rejection of the evaluated null hypotheses under the analytic conditions examined. The composite index should therefore be interpreted as a descriptive aggregation tool that summarizes relative performance patterns rather than as evidence of universal methodological superiority. Replication across additional datasets and modeling contexts would further clarify the stability and generalizability of these patterns.