



A Monte Carlo Simulation Study of the Psychometric Validation Strategy for the Residual Abandonment Index (RAI)

Lam Nam Phuong¹, Nguyen Minh Anh²

^{1,2} Students, Le Quy Don High School for the Gifted, Da Nang, Vietnam

ARTICLE INFO	ABSTRACT
Published Online: 16 October 2026	The Residual Abandonment Index (RAI) is a new multidimensional index proposed by the authors, who are high school students at Le Quy Don High School for the Gifted, Da Nang, Vietnam. RAI was developed by the authors to operationalize persistent abandonment-related experiences through five proposed dimensions. As a newly proposed index rather than an adaptation of an established scale, this study examines whether its proposed measurement structure can be statistically recovered under controlled conditions. A Monte Carlo simulation used a known five-factor population model with 25 items across 12 conditions crossing sample size ($N = 300, 600, 1,200$), primary loading (.50, .70), and interfactor correlation (.30, .60), with 200 replications per condition. Exploratory recovery used common-factor principal-axis factoring followed by an oblique Promax-type rotation, with confirmatory recovery additionally conducted on 20 replications per condition. EFA congruence, loading RMSE, internal consistency, Monte Carlo uncertainty, and CFA fit indices were evaluated. Mean congruence ranged from .866 to .997, with 11 of 12 conditions exceeding .90 and recovery rates ranging from .745 to 1.000. Confirmatory recovery produced mean CFI values above .997 with low RMSEA and SRMR across conditions. These findings indicate that the proposed five-factor measurement structure can be statistically recovered under the simulated conditions. However, the results do not constitute empirical validation of RAI in Vietnamese adolescents.
Corresponding Author: Lam Nam Phuong	
KEYWORDS: Academic collaboration, Payment facilitation, Author referrals, Publication fees, Editorial independence, Publication ethics.	

1. INTRODUCTION

1.1. Rationale

The Residual Abandonment Index (RAI) is a new multidimensional index proposed by the authors, two high school students at Le Quy Don High School for the Gifted, Da Nang, Vietnam. The index was developed independently by the authors and is not presented as an adaptation or inheritance of an established psychological scale.

The proposed RAI consists of five dimensions: Fear of Being Forgotten (F), Sensitivity to Being Ignored (I), Repeated Validation Seeking (V), Post-Validation Emptiness (E), and Anticipatory Abandonment (A). Because the index is newly proposed, the present study does not assume that this structure is valid. Instead, it first examines whether the proposed five-factor measurement structure can be statistically recovered under controlled simulated conditions. This distinction is important: a successful simulation does not establish that RAI is valid in a real population. Rather, it provides methodological evidence that, if the proposed

structure generated the data, the planned psychometric procedures could recover that structure with measurable accuracy.

1.2. Research Questions

- RQ1. How does sample size affect recovery of the proposed five-factor RAI structure?
- RQ2. How does primary item loading affect factor recovery and internal consistency?
- RQ3. How does interfactor correlation affect the distinguishability of the five dimensions?
- RQ4. Under which conditions is the proposed validation workflow sufficiently stable for subsequent empirical testing?

1.3. Expanded Literature Review and Methodological Positioning

The proposed RAI is positioned at the intersection of several established literatures rather than as a replacement for them. Attachment theory provides a developmental account of

expectations concerning relational availability; rejection-sensitivity research addresses the interpretation and response to possible rejection; and contingencies-of-selfworth theory explains how external approval can become linked to self-evaluation. These literatures provide convergent theoretical foundations for the proposed dimensions but do not, by themselves, establish the RAI as a validated construct. The present study therefore treats the five dimensions as a researcher-generated measurement hypothesis requiring systematic psychometric evaluation.

Methodologically, the simulation rationale follows evidence that factor recovery depends on characteristics of the measurement model rather than on sample size alone. MacCallum et al. (1999) showed that required sample size in factor analysis depends substantially on communalities and factor overdetermination. Wolf et al. (2013) similarly demonstrated through Monte Carlo analyses that sample-size requirements in latent-variable models vary with model characteristics such as loading magnitude, number of indicators, and factor structure. These findings support simulation-based planning when the goal is to understand recovery under explicitly defined population conditions rather than relying on a universal participant-to-item ratio.

For exploratory factor analysis, methodological choices concerning factor extraction, retention, rotation, and loading evaluation can materially affect conclusions. Because the proposed RAI dimensions are theoretically related, an oblique common-factor solution is more conceptually appropriate for the eventual empirical validation than an orthogonal component solution. The current simulation is consequently best understood as a first-stage stress test; subsequent simulation work should explicitly compare common-factor extraction and oblique rotation procedures before empirical data collection.

The literature therefore supports the central methodological logic of the present study: a researcher-developed index should be evaluated through an explicit theoretical model, transparent item and factor specifications, controlled simulation conditions, and empirical validation using real observations. The Monte Carlo study does not substitute for that empirical validation; it provides information about whether the proposed statistical recovery strategy is sufficiently robust to justify the next stage of investigation.

2. THEORETICAL BACKGROUND AND CONCEPTUAL MEASUREMENT MODEL

2.1. Conceptualizing Residual Abandonment-Related Vulnerability

The Residual Abandonment Index (RAI) is grounded in the proposition that experiences of relational absence, emotional unavailability, instability, or perceived abandonment may leave enduring socio-emotional tendencies even when an individual does not meet criteria for a psychiatric disorder.

The term residual is therefore used to describe persistent psychological patterns that may remain after an adverse relational experience rather than to imply a clinical syndrome. This distinction is important because relational adversity and its psychological consequences are not necessarily equivalent to psychopathology. Adolescents may exhibit heightened sensitivity to interpersonal cues, recurrent reassurance seeking, concern about losing relational significance, or expectations of future rejection without necessarily displaying a diagnosable mental disorder. The RAI is therefore conceptualized as a non-clinical, population-level research index rather than a diagnostic instrument.

2.2. Attachment Theory as a Theoretical Foundation

Attachment theory provides one of the principal theoretical foundations for understanding why early relational experiences may influence later expectations of interpersonal availability and rejection. Bowlby's attachment framework proposes that repeated interactions with caregivers contribute to internal working models concerning the self, others, and the expected availability of significant relationships. From this perspective, inconsistent, unavailable, or disrupted caregiving may contribute to expectations that important others are unreliable or that relational security is fragile. Such expectations may subsequently influence how ambiguous interpersonal events are interpreted. However, attachment theory alone does not justify treating RAI as equivalent to attachment anxiety. Attachment anxiety represents an established relational attachment construct, whereas RAI is designed to operationalize a more specific constellation of abandonment-related manifestations. This distinction makes discriminant validity between RAI and attachment anxiety a central empirical question.

2.3. Rejection Sensitivity and Interpretation of Interpersonal Cues

A second theoretical foundation is provided by research on rejection sensitivity. Rejection sensitivity concerns the tendency to anxiously expect, readily perceive, and strongly react to rejection. This framework is particularly relevant to the RAI dimension Sensitivity to Being Ignored (I). An adolescent with heightened interpersonal vigilance may assign greater emotional significance to ambiguous cues such as delayed communication, exclusion, reduced attention, or perceived interpersonal distance. Within the RAI framework, sensitivity to being ignored is not intended to reproduce the entire rejection-sensitivity construct. Instead, it focuses specifically on the interpersonal experience of being overlooked, forgotten, excluded, or insufficiently attended to as a possible manifestation of abandonment-related vulnerability.

2.4. External Validation and Contingent Self-Worth

A third theoretical component concerns the relationship between interpersonal approval and self-evaluation.

Contingencies of self-worth theory proposes that individuals may base their sense of personal value on obtaining success or approval in particular domains. When self-worth becomes strongly dependent on external evaluation, positive feedback may temporarily increase self-evaluation, whereas the absence or withdrawal of approval may produce psychological instability. This framework provides a theoretical basis for the RAI dimensions Repeated Validation Seeking (V) and Post-Validation Emptiness (E). Repeated validation seeking can be conceptualized as an interpersonal strategy through which an individual repeatedly obtains reassurance, affirmation, or evidence of relational significance. The proposed mechanism is: Perceived Insecurity → External Validation Seeking → Temporary Relief → Renewed Insecurity. The distinction between V and E is theoretically important. Validation seeking describes a behavioral or motivational response, whereas post-validation emptiness describes the subsequent affective state when external affirmation fails to produce enduring internal security.

2.5. Anticipatory Abandonment and Future-Oriented Interpersonal Expectations

Anticipatory Abandonment (A) represents a future-oriented component of abandonment-related vulnerability. Whereas Sensitivity to Being Ignored focuses primarily on responses to present or recent interpersonal cues, anticipatory abandonment concerns expectations that rejection, separation, or relational loss may occur in the future. The dimension therefore introduces a temporal component to the model: past relational adversity may contribute to present interpersonal vigilance, which may in turn be accompanied by expectations of future abandonment. Such expectations may be associated with heightened monitoring, pre-emptive withdrawal, or attempts to protect oneself from an expected future loss. The construct does not assume that the expected abandonment will actually occur. It concerns the expectation or anticipation of relational loss and the psychological consequences of that expectation.

2.6. Emotion Regulation as a Possible Mechanistic Link

The proposed RAI dimensions can also be situated within models of emotion regulation. Interpersonal experiences are closely connected to emotional appraisal and regulation, particularly during adolescence. Repeated validation seeking may function as an interpersonal emotion-regulation strategy. An individual experiencing uncertainty or distress may seek reassurance from others to reduce negative affect. If reassurance provides only temporary relief, the individual may subsequently return to a similar state of uncertainty, creating a repetitive regulatory cycle. This provides a theoretical bridge between the cognitive, affective, and behavioral components of RAI. The proposed index is therefore intended to capture how relational vulnerability

may be expressed psychologically after an adverse experience, rather than simply counting adverse events.

2.7. Adolescence as a Developmentally Relevant Context

Adolescence provides a theoretically relevant developmental context for investigating abandonment-related socioemotional processes. During adolescence, peer relationships become increasingly important, identity becomes more differentiated, and social evaluation may acquire greater psychological significance. Consequently, interpersonal cues such as exclusion, delayed responses, changes in friendship, or perceived loss of importance may become particularly salient. The Vietnamese context may further shape these processes through family structure, academic expectations, interpersonal norms, peer relationships, and digitally mediated communication. The purpose of this contextualization is not to claim that RAI is culturally unique to Vietnam. Rather, Vietnamese adolescents constitute an important population in which the proposed construct can be examined empirically and culturally contextualized.

2.8. The Five-Dimensional Structure of RAI

Taken together, the theoretical foundations above suggest that the five proposed dimensions represent related but potentially distinguishable manifestations of a broader abandonment-related vulnerability. The five dimensions are: F = Fear of Being Forgotten; I = Sensitivity to Being Ignored; V = Repeated Validation Seeking; E = Post-Validation Emptiness; A = Anticipatory Abandonment. The conceptual model can therefore be expressed as: Residual Abandonment-Related Vulnerability = $f(F, I, V, E, A)$. The five dimensions should not be treated as empirically established merely because they are theoretically plausible. Their separability, correlations, and higher-order organization remain empirical questions. This principle is particularly important because RAI is a researcher-developed index rather than an inherited or directly adapted established scale.

2.9. RAI in Relation to Established Constructs

The theoretical position of RAI can be summarized through its relationship with several established domains. Attachment theory provides a foundation for understanding relational expectations; attachment anxiety is a related but non-identical comparator; rejection sensitivity helps explain interpersonal vigilance and perceived neglect; contingencies of self-worth provide a basis for validation-dependent self-evaluation; emotion regulation offers a possible mechanism; and adverse relational experiences provide developmental context. Accordingly, RAI is not proposed as an isolated psychological phenomenon. It occupies an intersection among established theoretical domains while attempting to capture a particular configuration that is not assumed to be identical to any one of them.

2.10. Theoretical Propositions

The theoretical model generates several propositions for subsequent validation. First, the five RAI dimensions should exhibit meaningful positive associations because they represent related manifestations of a broader vulnerability. Second, they should nevertheless demonstrate sufficient empirical separability to justify multidimensional measurement. Third, RAI should correlate positively with attachment anxiety but should not be statistically indistinguishable from it. Fourth, RAI should show meaningful associations with approval-contingent self-worth. Fifth, RAI may provide incremental statistical information beyond attachment anxiety if its configuration captures processes not fully represented by attachment anxiety.

2.11. Implications for the Monte Carlo Simulation Study

The theoretical model provides the rationale for the simulation design. Because RAI is researcher-developed, the simulation does not attempt to demonstrate that the construct exists in a population. Instead, it evaluates whether a prespecified measurement structure can be statistically recovered under controlled conditions. The central methodological question is: Given a prespecified five-factor

RAI model, under what conditions can its structure be recovered reliably? Sample size, factor loading magnitude, and interfactor correlation are therefore treated as key simulation parameters. Stronger loadings should generally provide clearer indicators of their intended factors, whereas higher interfactor correlations may make dimensions more difficult to distinguish. The simulation is consequently a methodological stress test of the proposed measurement model rather than empirical validation of RAI in Vietnamese adolescents. Simulation Evidence $\neq$ Population Validity Evidence. Instead, simulation evidence concerns statistical recoverability, whereas empirical validation requires evidence from real participants. A successful simulation strengthens the methodological justification for empirical validation; an unsuccessful simulation indicates that the measurement model or recovery procedure requires refinement before empirical implementation.

The proposed RAI measurement model contains five correlated latent dimensions. Each dimension is represented by five observed items. The population model is intentionally specified before simulation so that the true parameters are known and can be compared with the parameters recovered by the analysis.

Population measurement equation

$$\mathbf{X} = \mathbf{A}\mathbf{F} + \boldsymbol{\varepsilon}$$

$\mathbf{X}$ = observed item responses; $\mathbf{A}$ = population loading matrix; $\mathbf{F}$ = five latent RAI dimensions; $\boldsymbol{\varepsilon}$ = residual error.

RAI composite score

$$\text{RAI} = (\mathbf{F} + \mathbf{I} + \mathbf{V} + \mathbf{E} + \mathbf{A}) / 5$$

Equal weighting is the primary scoring rule proposed for transparency; alternative weighting is treated as sensitivity analysis.

3. METHODS

3.1. Simulation Design

A Monte Carlo design was implemented with 12 experimental conditions. Sample size, primary factor loading,

and interfactor correlation were manipulated systematically. Each condition was independently replicated 200 times, yielding 2,400 synthetic datasets in total.

Table 1. Prespecified Monte Carlo simulation conditions.

N	Primary loading λ	Interfactor correlation r
300	.50	.30
300	.50	.60
300	.70	.30
300	.70	.60
600	.50	.30
600	.50	.60
600	.70	.30
600	.70	.60
1200	.50	.30
1200	.50	.60
1200	.70	.30
1200	.70	.60

3.2. Data-Generating Process

Five correlated latent variables were generated from a multivariate normal distribution. Each latent factor was connected to five items. The primary loading was set to either

.50 or .70, and residual variance was chosen as $1 - \lambda^2$ so that the continuous item variance was standardized before ordinal transformation.

Residual variance specification

$$\text{Var}(\epsilon_j) = 1 - \lambda_j^2$$

This specification produces standardized continuous indicators under the population model.

Continuous responses were discretized into five ordered categories using fixed thresholds, yielding synthetic Likert-type responses. No real participant response was used at any stage.

3.3. Monte Carlo Replication

For each condition c , the simulation was repeated $R=200$ times. Let $\theta_{c,r}$ denote a recovery statistic from replication r under condition c . The condition-level Monte Carlo estimate was calculated as the arithmetic mean across replications.

$$\text{Monte Carlo estimate } \hat{\theta}^{\text{MC},c} = (1/R) \sum_{r=1}^R \theta_{c,r}$$

$R = 200$ replications per condition.

Monte Carlo standard error

$$\text{MCSE}(\hat{\theta}^{\text{MC},c}) = \text{SD}(\theta_{c,r}) / \sqrt{R}$$

Lower MCSE indicates greater numerical stability of the Monte Carlo estimate.

3.4. Factor-Recovery Analysis

Exploratory recovery used common-factor principal-axis factoring (PAF) on the standardized item correlation matrix, followed by an oblique Promax-type rotation. Five factors were retained because the population model contained five correlated latent dimensions. This approach was selected to

align the recovery procedure with the latent-variable interpretation of RAI and the expectation that its dimensions are correlated. Factor-pattern similarity was evaluated against the known population loading matrix after optimal factor matching and sign alignment.

Factor congruence

$$\phi = \sum_i (a_i b_i) / \sqrt{[\sum_i (a_i^2) \sum_i (b_i^2)]}$$

Tucker-type congruence was evaluated after matching estimated factors to the known population factors.

Loading recovery error

$$\text{RMSE}_\lambda = \sqrt{[(1/p) \sum_j (|\hat{\lambda}_j| - |\lambda_j|)^2]}$$

Lower RMSE indicates closer recovery of the population loading magnitudes.

3.5. Reliability Analysis

3.5.1. Statistical Methodology for Monte Carlo Evaluation

The simulation was designed as a fully crossed Monte Carlo experiment in which population sample size (N), standardized factor loading magnitude (λ), and interfactor correlation (ρ) were manipulated systematically. These parameters were selected because factor recovery is influenced by the amount of information available in the sample, the strength of the relationship between observed indicators and their latent factors, and the degree of overlap among correlated factors. Prior simulation work likewise demonstrates that sample size and loading magnitude can materially affect parameter recovery and solution propriety, making simple sample-size rules insufficient as a universal

basis for psychometric planning (MacCallum et al., 1999; Wolf et al., 2013).

Each simulation condition was replicated $R = 200$ times. For every replication, latent factor scores were generated from a multivariate normal distribution with unit factor variances and the prespecified interfactor correlation matrix. Observed item responses were then generated from the common-factor model $X = \Lambda F + \epsilon$, with residual variances determined from the specified loading magnitude. This procedure ensures that the population parameters are known exactly, allowing estimated quantities to be compared directly with their generating values.

Three complementary recovery criteria were evaluated. First, factor-pattern similarity was summarized using Tucker's congruence coefficient, which evaluates similarity between

estimated and population loading patterns. Second, loading recovery was evaluated using root mean square error (RMSE), with smaller values indicating closer recovery of the population loading structure. Third, internal consistency was summarized using Cronbach's alpha as a descriptive reliability indicator. Alpha was not treated as evidence of structural validity because internal consistency does not by itself establish dimensionality.

Monte Carlo uncertainty was quantified using the empirical standard deviation of each statistic across replications and the Monte Carlo standard error (MCSE). For a parameter estimate θ , the Monte Carlo mean was calculated as the average estimate across replications, whereas MCSE was calculated as the empirical standard deviation divided by the square root of R. Reporting MCSE distinguishes simulation noise from substantive differences between experimental conditions.

The simulation results were interpreted comparatively rather than against a single universal cutoff. Because the upgraded recovery procedure uses common-factor extraction with oblique rotation, the resulting congruence coefficients are interpreted as evidence about recovery of the prespecified latent pattern under the simulated conditions. A separate confirmatory recovery analysis used the known simple factor pattern and estimated factor loadings and factor correlations; its fit indices are therefore treated as recovery diagnostics rather than as empirical evidence that the RAI measurement model fits real participants.

Internal consistency was assessed using Cronbach's alpha. In the simulation, alpha describes the internal consistency of the synthetic indicators and should not be interpreted as evidence of reliability in human participants.

Cronbach's alpha

$$\alpha = k/(k-1) \times [1 - \Sigma \sigma^2_{\text{item}} / \sigma^2_{\text{total}}]$$

k = number of items; σ^2_{item} = item variance; σ^2_{total} = variance of the summed score.

4. RESULTS

All values in this section are simulated results. The purpose of the numerical results is to quantify the performance of the

proposed validation workflow under known synthetic conditions.

Table 2. Monte Carlo parameter-recovery results across all 12 simulation conditions.

SIMULATED/ILLUSTRATIVE.

Historical PCA analysis from the earlier draft; superseded by the common-factor EFA results in Table 2A.	λ	r	Mean ϕ	SD ϕ	95% empirical range	MCSE	RMSE λ	Mean α	$\phi \geq .85$
300	0.50	0.30	0.639	0.041	0.570–0.726	0.0029	0.283	0.717	0.0%
300	0.50	0.60	0.577	0.045	0.498–0.658	0.0032	0.288	0.811	0.0%
300	0.70	0.30	0.656	0.043	0.585–0.748	0.0030	0.359	0.847	0.0%
300	0.70	0.60	0.646	0.040	0.571–0.721	0.0029	0.363	0.911	0.0%
600	0.50	0.30	0.654	0.041	0.584–0.733	0.0029	0.279	0.716	0.0%
600	0.50	0.60	0.622	0.040	0.557–0.693	0.0028	0.284	0.813	0.0%
600	0.70	0.30	0.658	0.039	0.586–0.730	0.0028	0.357	0.849	0.0%
600	0.70	0.60	0.655	0.041	0.584–0.729	0.0029	0.362	0.912	0.0%

“A Monte Carlo Simulation Study of the Psychometric Validation Strategy for the Residual Abandonment Index (RAI)”

1200	0.50	0.30	0.652	0.040	0.588– 0.728	0.0028	0.278	0.720	0.0%
1200	0.50	0.60	0.640	0.041	0.564– 0.717	0.0029	0.283	0.813	0.0%
1200	0.70	0.30	0.657	0.042	0.581– 0.738	0.0030	0.354	0.848	0.0%
1200	0.70	0.60	0.660	0.040	0.586– 0.729	0.0029	0.362	0.912	0.0%

The strongest EFA recovery was observed under N=1200, $\lambda=0.70$, and $r=0.30$, with mean congruence $\phi=0.997$. The most challenging condition was N=300, $\lambda=0.50$, and $r=0.60$, with mean congruence $\phi=0.866$. All conditions produced mean congruence above .85, although the N=300, $\lambda=.50$, $r=.60$ condition showed greater variability and a lower recovery rate of .745.

Across conditions, common-factor EFA with oblique rotation produced high structural recovery. Higher factor correlations were associated with lower congruence, especially at N=300 and $\lambda=.50$, while larger samples and stronger loadings reduced variability and increased the stability of recovery.

Table 2A. Upgraded common-factor EFA and confirmatory-recovery results across all 12 simulation conditions.

N	λ	ρ	Mean ϕ	SD ϕ	2.5– 97.5%	RMSE λ	Mean α	$\phi \geq .85$	CFI	RMSEA	SRMR
300	0.50	0.30	0.9500	0.0121	0.9262– 0.9671	0.1966	0.6243	0.995	0.9987	0.0014	0.0504
300	0.50	0.60	0.8656	0.0504	0.7154– 0.9215	0.2046	0.6225	0.745	0.9983	0.0019	0.0494
300	0.70	0.30	0.9890	0.0019	0.9851– 0.9921	0.2199	0.8260	1.000	0.9991	0.0021	0.0433
300	0.70	0.60	0.9776	0.0043	0.9681– 0.9843	0.2496	0.8263	1.000	0.9988	0.0033	0.0413
600	0.50	0.30	0.9762	0.0043	0.9672– 0.9840	0.1980	0.6236	1.000	0.9992	0.0006	0.0350
600	0.50	0.60	0.9413	0.0133	0.9081– 0.9615	0.2060	0.6239	1.000	0.9977	0.0025	0.0355
600	0.70	0.30	0.9947	0.0008	0.9930– 0.9963	0.2197	0.8273	1.000	0.9984	0.0032	0.0312
600	0.70	0.60	0.9889	0.0021	0.9846– 0.9924	0.2499	0.8272	1.000	0.9994	0.0016	0.0296
1200	0.50	0.30	0.9882	0.0020	0.9840– 0.9917	0.1985	0.6258	1.000	0.9993	0.0010	0.0250
1200	0.50	0.60	0.9718	0.0051	0.9594– 0.9801	0.2065	0.6249	1.000	0.9998	0.0005	0.0247
1200	0.70	0.30	0.9973	0.0004	0.9964– 0.9982	0.2195	0.8275	1.000	0.9998	0.0011	0.0219
1200	0.70	0.60	0.9945	0.0009	0.9925– 0.9961	0.2500	0.8280	1.000	0.9999	0.0006	0.0205

Table 3. Marginal EFA recovery by sample size and loading strength. SIMULATED.

N	Primary loading	Mean congruence across factor correlation conditions
300	0.50	0.608
300	0.70	0.651

600	0.50	0.638
600	0.70	0.657
1200	0.50	0.646
1200	0.70	0.659

Table 4. Simulated internal consistency under the upgraded common-factor model. SIMULATED.

Primary loading	Mean α across conditions	Range of condition means
0.50	0.765	0.716–0.813
0.70	0.880	0.847–0.912

4.1. Confirmatory Recovery Analysis

The confirmatory component was evaluated on 20 replications per condition (every tenth replication of the 200 EFA replications). The CFA used the known simple five-factor pattern, with five indicators per factor and freely estimated factor correlations. Because the data-generating model and the fitted model share the same factor pattern, the resulting fit indices are interpreted as confirmatory recovery diagnostics under correct specification rather than as empirical evidence of construct validity. Across conditions, mean CFI ranged from .9977 to .9999, mean RMSEA from .0005 to .0033, and mean SRMR from .0205 to .0504.

5. INTERPRETATION

5.1. Evidence for Statistical Recoverability

Because the RAI structure was specified by the researcher before simulation, successful recovery can be interpreted as evidence that the proposed statistical workflow is capable of detecting the hypothesized structure under the simulated conditions. This is a methodological proof-of-recoverability rather than proof of psychological validity.

5.2. Why This Matters for a Researcher-Developed Index

RAI is intentionally presented as a researcher-developed index. The simulation is therefore not an attempt to demonstrate that RAI inherits validity from an established scale. Instead, it provides a controlled test of whether the measurement architecture proposed by the researcher can generate and recover a coherent multidimensional structure. The logical sequence is therefore:

- Conceptual specification of the RAI construct.
- Generation of candidate indicators.
- Specification of a five-factor population model.
- Monte Carlo testing of structural recoverability.
- Empirical administration to real participants.
- EFA/CFA and external validity testing on observed data.
- Revision or rejection of the proposed model if empirical evidence does not support it.

5.3. Discussion of the Monte Carlo Findings

The upgraded simulation produced strong structural recovery under the common-factor EFA procedure. Across all 12

conditions, mean factor congruence ranged from .8656 to .9973. The lowest mean occurred at $N=300$, $\lambda=.50$, $\rho=.60$, whereas the highest occurred at $N=1200$, $\lambda=.70$, $\rho=.30$. Eleven of the twelve conditions produced mean congruence above .90, and the recovery rate for $\phi \geq .85$ ranged from .745 in the most challenging condition to 1.000 in the remaining conditions. This pattern indicates that the prespecified five-factor population structure is statistically recoverable under the simulated measurement conditions when a common-factor method and oblique rotation are used.

The contrast with the earlier PCA-based analysis is methodologically informative. The present results show substantially stronger recovery when the extraction method is aligned with the latent-factor model. This demonstrates why PCA should not be treated as interchangeable with common-factor EFA in a psychometric validation study. The improvement in congruence is therefore evidence about the recovery procedure, not evidence that the underlying psychological construct has already been established.

Internal consistency followed the expected loading pattern. Mean alpha was approximately .62–.63 under $\lambda=.50$ and approximately .83 under $\lambda=.70$. Thus, stronger population loadings increased internal consistency, while factor correlation had relatively little effect on alpha. These values also reinforce the importance of not treating alpha as a standalone validity criterion: structural recovery can be strong even when internal consistency is only moderate under lower-loading conditions.

Sample size primarily influenced the stability of recovery rather than producing a simple threshold effect. At $N=300$, the combination of lower loading and higher interfactor correlation generated the greatest variability and the only recovery rate below 1.00. Increasing N to 600 and 1,200 substantially reduced the dispersion of congruence estimates and produced uniformly high recovery. This pattern is consistent with simulation literature showing that factor recovery requirements depend jointly on sample size, communalities/loading strength, model complexity, and factor structure (MacCallum et al., 1999; Wolf et al., 2013).

The results support proceeding to empirical validation, but they do not eliminate the need for further simulation work.

A stronger next step would vary additional realistic features such as unequal loadings, cross-loadings, non-normal indicators, ordinal response categories, missing data, and deliberate model misspecification. A CFA simulation should also evaluate parameter bias, confidence-interval coverage, convergence, improper solutions, and model-fit recovery under misspecified as well as correctly specified models. These extensions would provide a more stringent test of the proposed RAI measurement strategy.

5.4. Integrated Interpretation

Taken together, the upgraded simulation provides methodological evidence that the prespecified five-factor RAI model can be recovered with high similarity under the tested conditions when common-factor extraction and oblique rotation are used. The confirmatory recovery analysis likewise produced consistently high CFI and low RMSEA/SRMR, although these values must be interpreted as diagnostics under a correctly specified synthetic model rather

than as empirical fit evidence. The appropriate conclusion is therefore that the proposed measurement model is computationally recoverable under the simulated assumptions, not that RAI has already been validated in a population.

This interpretation is consistent with the status of RAI as a researcher-developed index. The scientific contribution at this stage lies in making the construct definition, measurement assumptions, simulation parameters, and recovery criteria explicit and reproducible. Empirical validity remains a separate question that must be addressed using real participants and prespecified psychometric analyses.

6. Sensitivity Analysis

The simulation conditions provide a sensitivity analysis of the proposed measurement model. In particular, the comparison between $\lambda=.50$ and $\lambda=.70$ evaluates how dependent structural recovery is on indicator quality, while $r=.30$ versus $r=.60$ evaluates how factor overlap influences recovery.

Recovery-rate estimator

$$\hat{P}(\phi \geq .85) = (1/R) \sum I(\phi_r \geq .85)$$

This is a descriptive operating-characteristic measure, not a psychometric validity criterion.

7. WHAT THE SIMULATION ESTABLISHES

- It establishes the statistical behavior of the specified recovery procedure under the simulated population model.
- It quantifies the sensitivity of factor recovery to sample size.
- It quantifies the sensitivity of factor recovery to item loading strength.
- It demonstrates how interfactor correlation affects structural recovery.
- It provides methodological guidance for planning the subsequent empirical study.

8. WHAT THE SIMULATION DOES NOT ESTABLISH

- It does not establish that RAI is a valid psychological construct in Vietnamese adolescents.
- It does not establish discriminant validity from attachment anxiety, rejection sensitivity, or approval-contingent self-worth in real participants.
- It does not establish prevalence or population-level RAI scores.
- It does not establish empirical criterion or incremental validity.
- It does not replace ethical human-participant research.

9. FUTURE EMPIRICAL VALIDATION

The next stage should be a separate empirical study involving real participants. The empirical study should retain the conceptual model only as a testable hypothesis. Candidate

items should undergo expert content review and cognitive testing before field administration. EFA and CFA should be conducted using prespecified procedures, preferably with independent or split samples. Reliability, convergent validity, discriminant validity, and incremental validity should then be evaluated against appropriate established comparator measures.

10. LIMITATIONS

- The simulation is conditional on the assumed data-generating model.
- A correctly recovered model may still be theoretically incorrect in the real population.
- The present simulation emphasizes structural recovery rather than the full range of psychometric threats.
- Ordinal-response behavior, missingness, cross-loadings, differential item functioning, and model misspecification can be investigated in additional simulation scenarios.
- Monte Carlo replication error remains finite; larger replication counts can provide more precise operating characteristic estimates.

11. CONCLUSION

The present Monte Carlo study treats RAI as a genuinely researcher-developed index and evaluates its proposed measurement architecture without borrowing empirical validity from an inherited instrument. The simulation demonstrates whether the specified psychometric workflow can recover the proposed five-factor structure under controlled synthetic conditions.

The appropriate scientific claim is therefore not that RAI has already been validated in Vietnamese students. Rather, the claim is that the proposed RAI measurement model has been subjected to a controlled simulation of parameter recovery, providing methodological evidence to guide subsequent empirical validation.

12. DATA, CODE, AND ETHICS STATEMENT

All observations in the present simulation are synthetic. No human participant data were used. The simulation seed, data-generating parameters, replication count, analysis procedures, and generated summary statistics should be archived with the manuscript. A subsequent empirical study involving minors must obtain appropriate ethical approval and age-appropriate consent/assent.

13. SIMULATION STUDY — SYNTHETIC DATA ONLY

The Residual Abandonment Index (RAI) is a researcher-developed index proposed by the authors. No human participant validation is claimed; all numerical results are based on synthetic data generated through Monte Carlo simulation.

REFERENCES

1. Fabrigar, L. R., & Wegener, D. T. (2012). Exploratory Factor Analysis. Oxford University Press.
2. Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, 76, 537–549. <https://doi.org/10.1007/s11336-011-9218-4>
3. MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99. <https://doi.org/10.1037/1082-989X.4.1.84>
4. Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913–934. <https://doi.org/10.1177/0013164413495237>
5. Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human–Computer Interaction*, 32(1), 2–17. <https://doi.org/10.1080/10447318.2015.1087664>
6. Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8
7. The terms 'Monte Carlo simulation', 'synthetic data', 'known population parameters', 'parameter recovery', and 'operating characteristics' should be used consistently. The simulated results must not be described as observations from Vietnamese high-school students.

Appendix A. Simulation Parameter Specification Table A1. Complete simulation specification.

Parameter	Value	Specification
Latent dimensions	5	F, I, V, E, A
Items per dimension	5	25 items total
Response categories	5	Ordinal Likert-type
Sample sizes	300; 600; 1,200	12 factorial conditions
Primary loadings	.50; .70	Manipulated
Interfactor correlations	.30; .60	Manipulated
Replications per condition	200	2,400 datasets total
Factor recovery criterion	.85	Descriptive only
Random seed	20261006	Fixed for reproducibility

Appendix B. Formula Summary

1. Measurement model

$$X = \Lambda F + \varepsilon$$

2. Residual variance

$$\text{Var}(\varepsilon_j) = 1 - \lambda_j^2$$

3. RAI equal-weight score

$$\text{RAI} = (F + I + V + E + A) / 5$$

4. Monte Carlo

$$\text{mean } \hat{\theta}^{\text{MC}} = (1/R) \sum \theta_r$$

5. Monte Carlo standard error

$$\text{MCSE} = \text{SD}(\theta_r) / \sqrt{R}$$

6. Factor congruence $\phi = \Sigma(a_i b_i) / \sqrt{[\Sigma(a_i^2) \Sigma(b_i^2)]}$

7. Loading RMSE

$$\text{RMSE}_\lambda = \sqrt{[(1/p) \Sigma(|\hat{\lambda}_j| - |\lambda_j|)^2]}$$

8. Cronbach's alpha $\alpha = k/(k-1) \times [1 - \Sigma \sigma^2_{\text{item}} / \sigma^2_{\text{total}}]$

9. Recovery rate

$$P^*(\phi \geq .85) = (1/R) \sum I(\phi_r \geq .85)$$

Appendix C. Interpretation Key

Table C1. Interpretation of simulation outputs.

Statistic	Interpretation
Mean congruence	Average similarity between estimated and known factor structures.
MCSE	Numerical uncertainty caused by finite Monte Carlo replication.
Loading RMSE	Average magnitude of loading-recovery error.
Recovery rate	Proportion of simulated replications meeting the descriptive $\phi \geq .85$ criterion.
Mean α	Average internal consistency across synthetic datasets.