The False Allure of Causal Indicator Models

Frank D. Mann, PhD*[a] & Kristian E. Markon, PhD[b]

[a] Department of Medicine, Renaissance School of Medicine at Stony Brook University

[b] Department of Psychology, University of Minnesota, Twin Cities

***Correspondence:** Dr. Frank D. Mann, PhD, Department of Medicine, 101 Nichols Rd., Health Sciences Center, #3-071, Renaissance School of Medicine at Stony Brook University, Stony Brook, NY, 11794, USA. Email: frank.mann@stonybrookmedicine.edu.

ABSTRACT

Advocates of “causal” indicator measurement have argued that adverse childhood experiences should *form* a latent adversity construct that is then linked to health outcomes. We contend this proposal rests on a mistaken premise: common factor models merely decompose covariance and do not assume that an unobserved entity causes its indicators; causal interpretations require design features, not arrow direction. Crucially, reflective models already realize the representational goal formative advocates seek, quantifying information loss through residual variances. Besides, pure formative blocks are not locally identified and must compete to explain the same shared variance, inflating standard errors. In contrast, reflective models are locally identified and leverage the interdependence of adverse childhood experiences. Taken together, these points reveal formative factors are either disturbance-free composites or identified by their reflective anchors, whereas reflective models provide a fully identified and testable measurement model.

The False Allure of "Causal" Indicator Models

Researchers have long debated the relative merits of formative causal indicator versus reflective measurement models[1-4]. Notably, over the past decade, causal indicator models have gained traction across disciplines where constructs are viewed as composites of distinct causal influences. For instance, a recent recommendation encourages developmental scientists to model adverse childhood experiences (ACEs) with causal indicators, on the grounds that this approach can yield more precise stress profiles and enhance translational utility for both patients and clinicians[5]. At first glance, the proposal seems reasonable: specify models in which discrete stressors create adversity, then relate that construct to downstream health. After all, that is what theory dictates. However, a closer look reveals that common factor models are often misconstrued as presuming the latent factor is a cause of its indicators, when in fact the model can be viewed more neutrally, understood as simply a covariance decomposition, without any assumption that an unobserved entity causes its indicators. Moreover, contrary to what might be assumed, reflective models instantiate representational measurement — arguably the central compelling feature of formative approaches, and the type of measurement being advocated for by proponents of causal indicator models — better than formative models.

Arrow Illusions: Covariance Without Causation

Although path diagrams have become the lingua franca of structural equation modeling, their graphical conventions — especially the unidirectional arrows streaming from a latent factor toward its indicators — inadvertently nurture the notion that common factor models presume a literal causal flow from an unobserved construct to measured indicators. Indeed, as previously summarized[6], reflective and formative measurement models are often described using causal language[5,7-13]. However, the arrows in a path diagram of a common factor model merely encode a set of covariance restrictions: they remind the reader which elements of the loading matrix are freely estimated and which remain fixed, nothing more. Render the same model in compact matrix form and the suggestive imagery of directional influence disappears. The constraints now appear as algebraic relations among variances and covariances, making it clear the model is silent about causation.

Formally, the common factor model specifies the vector of observed variables (e.g., ACE items) as: $x = \Lambda f + \varepsilon,\ E[f] = 0,\ Var(f) = \Phi,\ Var(\varepsilon) = \Psi,\ Cov(f, \varepsilon) = 0$, which implies the population covariance matrix:

$$\sum\nolimits_{(p\times p)} = \Lambda\Phi\Lambda^{T} + \Psi \qquad (1)$$

Where x is the p-dimensional vector of observed variables; Λ is the p × k loading matrix whose ($\lambda_{i,j}$) entry represents how strongly observed variable x_i associates with latent factor f_i; f is the k-dimensional vector of latent factors, which are unobserved variables that summarize the covariation among the x's; ε is the p-vector of unique (error or disturbance) terms, capturing variance in each x_i not shared with the factors; $E[f] = 0$ sets the factor means to zero for identification and convenience, $Var(f) = \Phi$ specifies the covariance matrix of latent factors, $Var(\varepsilon) = \Psi$ is the matrix of unique variances, one per observed variable, and $Cov(f, \varepsilon) = 0$ states the latent factors and unique errors are uncorrelated, ensuring common and unique variances are statistically separate (but these error covariances can be allowed to be free to some extent in some models).

Under this specification the arrows in a path diagram merely encode the algebraic factorization $\sum$. The arrows do not test or assume that variation flows from the latent factors to the indicators in a temporal or causal sense. In fact, for every identified common factor model there exists an observationally equivalent formative model in which the arrows are reversed and the same $\sum$ is reproduced[14]. What must hold is only the linear constraints embodied in Λ and Φ describe the second-moment structure of the data. For example, one can algebraically re-express the latent variable formatively as predicted by the indicators:

$$f = \gamma^T x + \zeta \qquad (2)$$

where f = the 1×1 (or more generally r×1) vector of latent factor scores, γ^T = the transpose of the p×1 vector of factor-score weights, x = the p×1 vector of observed indicators (e.g., ACE items), and ζ = the scalar (or r×1 vector) of residual error, where var(ζ) is chosen so the total $Var(f)$ matches Φ (thus, ζ captures whatever part of the latent factor cannot be linearly predicted from x), with weights equal to:

$$\gamma = (\Lambda^T \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Lambda^T \Psi^{-1} \qquad (3)$$

Substituting this formative expression back into the measurement equation generates the same $\sum$—hence the models are observationally equivalent. There is a brief R tutorial in Appendix A demonstrating this. However, notice that γ is derived from the parameter estimates of the reflective model (Λ, Ψ, Φ) and not independently identified from Σ. In fact, γ is the weight matrix for a traditional regression-based factor score (a.k.a. Thomson or Thurstone score)[15,16]:

$$\hat{f}_{Th} = \Phi \Lambda^T \textstyle\sum^{-1} x = (\Lambda^T \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Lambda^T \Psi^{-1} x \qquad (4)$$

Indeed, factor scores are a formative re-expression of reflective parameters. Along similar lines, the ubiquitous unit-weighted sum score — often touted as a simple formative composite — is equivalent to a factor score from a highly constrained reflective model in which all loadings and residuals are fixed to one and zero, respectively[17]. Thus, what is presumed to be the simplest formative composite has a reflective counterpart.

So, are factor scores formative or reflective? In practice, they are both: a formative score whose weights inherit their meaning from a reflective structure, and a reflective model whose estimated scores are derived formatively. In this way, the formative-versus-reflective "cause-versus-effect" debate begins to dissolve, revealing two sides of the same covariance-analytic coin. Both the reflective and its mirror-image formative model are simply alternative, observationally equivalent factorizations of $\sum$, so any lingering sense of causal direction comes from us — not from the model.

Symmetry Shattered: The Reflective–Formative Divide

Extending the same notation to a model where arrows point from "causal" indicators to latent factor(s), the model-implied covariance is:

$$\textstyle\sum_{(y\times y)} = \Lambda(\Gamma \sum_{\xi\xi} \Gamma^T + \Phi)\Lambda^T + \Psi \qquad (5)$$

Where Γ = k × q matrix of regression coefficients that quantify the associations of the q-dimensional vector of exogenous variables ξ on the k-dimensional vector of latent factors, where each element represents the expected change in the latent factor given a one-unit increase in the exogenous variable ξ_j, when the other exogenous variables equal zero, and $\sum_{\xi\xi}$ is the $q \times q$ covariance matrix of the exogenous variables, which is known and provides the second-moment information that, together with Γ, links the exogenous variables to variability in the factor(s).

Two facts immediately follow. First, setting the structural paths to zero ($\Gamma = 0$) collapses the model back to the familiar reflective model (equation 1), leaving identification fully intact because the factor remains anchored by its reflective loadings, and its variance remains estimable through Φ. Second, setting the loadings to zero ($\Lambda = 0$) while leaving the exogenous "causal" block in place produces an empty factor that never reaches the data: its variance and regression parameters are not identified. The analyst must either delete the factor (yielding a trivial independence or saturated model) or impose ad-hoc constraints — e.g., setting parameters by fiat — that technically identify the model yet leave the latent variable untethered. This asymmetry underscores a weakness of “causal” indicator models: remove the formative block of exogenous variables (i.e., the “causal indicators”) and the measurement model remains fully identified. However, sever the reflective block and the entire measurement edifice collapses because the latent variable becomes under-identified[14,18-20].

Hence, what cannot be done in the mirror-image, pure formative specification of a reflective model is to estimate γ (equation 3) directly from $\sum$. The traffic of recoverability runs one way. A formative composite can be algebraically derived from a reflective solution — yet a “pure” formative block cannot return the favor by recreating the reflective loadings from the indicators alone. Until additional empirical anchors are added, the formative weights remain indeterminate and cannot be recovered from the data. Thus, you can “flip” the reflective model into a formative equation to yield observed scores, but you cannot identify the loadings for those scores independently of reflective parameters.

This asymmetry exposes an implicit hierarchy: reflective models are identifiable, testable, and capable of generating formative analogues, whereas formative models remain dependent on reflective structure for identification they do not independently supply. Crucially, causality lies outside the statistical model and must be justified on separate substantive methodological grounds. Thus, while imbuing directional arrows with causal inference might be pedagogically convenient, the validity of an applied common factor model hinges on its ability to reproduce the observed covariance matrix, not on a metaphysical claim that the latent variable causes variation in the indicators.

Composite by Fiat: The Day the Latent Went Missing

One route to rescue a pure formative block from being under-identified is to fix all formative weights to theory-derived constants (usually 1) and set the disturbance variance $\mathrm{Var}(\zeta)$

= 0. In that case, the factor is no longer a latent construct inferred from the data but rather a defined composite[12] — a sum (or weighted sum) of its indicators. Because we have removed the stochastic disturbance, the composite's variance is simply:

$$Var(\eta) = \gamma^T \Sigma_{xx} \gamma \tag{6}$$

which is a direct algebraic function of the observed indicator covariances and the fixed weights. No reflective anchors to downstream outcomes are needed in this case because identification is achieved by assumption rather than empirically. This composite model no longer estimates a latent variance or residual — it computes a summary score whose statistical properties are wholly determined by Σ_{xx} and the researcher's a priori γ.

This defined composite[12] approach might be appropriate when a researcher's substantive theory truly posits that the construct is nothing more than a deterministic combination of its parts, and when measurement error — or, equivalently, the variability in measurements across samples — is negligible or unimportant. In such settings, however, researchers forfeit the ability to evaluate model fit, assess measurement invariance, or model residuals. Thus, defining an error-free composite is best reserved for cases where theory demands a deterministic index rather than a latent variable, and where there is no interest in making inferences about components of measurements that are constant across samples or occasions.

In contrast, adverse childhood experiences do not constitute a scenario where an ignorable-error index is warranted. ACEs are heterogeneous and interdependent events[21-23]. Treating them as a unit-weighted, error-free sum tacitly assumes each experience has identical impact and measurement precision. A fixed γ=1 composite, therefore, lacks the flexibility to capture differential associations: it forces equal weight on each item, preventing the data from revealing which ACEs most potently measure adversity and predict outcomes. Consequently, a defined composite not only misrepresents theoretical nuance but also precludes insights into the distinct pathways linking adversities to long-term outcomes. Alternatively, if the analyst fixes the weights but wishes to freely estimate the latent (disturbance) variance, as advocated by proponents of “causal” indicator models[5], the model remains under-identified unless it is anchored by (a) at least two reflective, error-permitting, indicators, (b) one perfectly reliable (error-free) reflective indicator, or (c) some other pair of non-redundant exogenous relations (e.g., multiple structural paths or group/mean constraints) are added to anchor it.

Scaffolding Required: Parsimony Not Included

“Causal” indicator models are more complex than their reflective counterparts — not because they necessarily estimate more free parameters, but because they demand extra modelling layers merely to become identified and interpretable. A reflective model is identified by the covariances the indicators share. Apart from a single scaling constraint, all loadings, residuals, and factor covariances can be recovered directly from that indicator block, and the model can be judged with the familiar battery of invariance tests.

By contrast, a formative block supplies regression equations yet do not identify the latent variance, leaving the factor indeterminate until the analyst imposes additional constraints. As previously noted, one can fix formative weights, constrain the weights to sum to one while fixing the disturbance variance, or — more commonly — append at least two reflective outcomes[12,14,24,25]. Those anchors introduce new paths, residual correlations, and a disturbance variance that must be estimated jointly with the formative weights. The raw parameter count may be similar to, smaller than, or larger than the reflective analogue, but the complexity and compelled assumptions are invariably greater: researchers must defend choice of downstream markers, justify residual covariance specifications among indicators, and craft ad hoc diagnostics because the usual global-fit statistics (χ^2, CFI, RMSEA) remain computable for the whole structural equation model yet lack validated benchmarks for the formative block itself[11,26]. Formative weights are also conditional on the chosen reflective anchors and can shift — or even reverse sign — when those anchors change, especially under multicollinearity or small samples. Thus, even when the tally of free parameters is comparable, formative models impose greater identification hurdles, diagnostic uncertainty, and interpretational ambiguity than a reflective model.

Competitive Inferential Interference

Adverse childhood experiences tend to cluster empirically[21-23] — for example, household substance use with domestic violence, parental separation with neglect, physical abuse with emotional abuse — and that interdependence usually does not interfere with measurement parameter inferences in a reflective model (e.g., the standard errors of loadings). However, in a formative model correlated indicators must compete to explain the same shared variance, inflating standard errors. A sample with a different co-occurrence profile will also yield a different weight pattern. Although reflective loadings are also sample-specific, high inter-item correlations do not jeopardize parameter inferences or conceptual clarity — despite the potential for factor collapse in the extreme — because reflective loadings do not partition variance among themselves. Moreover, interpreting a Γ-coefficient as the expected change in adversity produced by a one-unit increase in ACE j when all the other ACEs equal 0 is awkward because the standard "all-else-equal" regression interpretation does not map cleanly onto the reality of clustered childhood adversities.

Modeling Sleight of Hand

As discussed above, pure "causal" indicator models are under-identified in isolation and inestimable. To make such models estimable, at least two auxiliary constraints are required—often two reflective indicators[24]. In a stress framework, those reflective indicators are variables such as HPA-axis dysregulation, neurocognitive impairment, allostatic load, or internalizing symptoms. Once they enter the model, the latent factor is shaped by the shared variance of those variables[27] rather than by variation in the ACE items only, which then act as exogenous regressors. In a "causal" indicator model, labeling the factor "latent adversity" would be a category mistake: if memory and executive function serve as anchors, the construct should be labeled "latent cognition"; if BMI and waist circumference are used, "latent adiposity", and so on. The ACE block itself lacks a standalone measurement component and thus cannot define the measurand. Indeed, "a latent variable with causal indicators cannot possibly derive its meaning from the causal indicators themselves"[28].

Alternative identification constraints, like fixing the latent variance or anchoring one x-to-f path at unity — can, in principle, dispense with reflective indicators as scaling devices. The hitch, however, is that those fixes still leave all remaining ACE weights, the latent mean, and the disturbance variance for the latent factor to be estimated conditional on the covariance matrix of the reflective block. In other words, even if one were to fix the scale in this way, the reflective outcomes that are hypothesized to be downstream of ACEs continue to exert leverage on the pattern of formative coefficients through the likelihood function. Therefore, the downstream markers do more than merely supply a metric; they co-determine the latent factor.

Enter the Covariate Cavalry

Including an exogenous block is warranted when theory or prior evidence suggests that variables outside the focal latent system — such as age, sex, race/ethnicity, or intervention assignment — may account for variability in latent constructs or their observed indicators. By regressing the latent factor(s) on these covariates (and, where justified, on selected indicators), the analyst can test hypotheses about demographic or experimental group differences in latent variables, partial out nuisance variance that would otherwise inflate unexplained variance and reduce bias in structural paths among endogenous variables. The exogenous block should be specified only when the covariates are temporally and conceptually prior, are measured without serious error, and are not themselves mediators, colliders, or consequences of the latent constructs under study. Importantly, adding such a block does not convert the measurement portion (Λ & Ψ) into a causal model. The loadings that map latent factors onto reflective indicators remain covariance constraints, so the introduction of exogenous regressors should not be mistaken for evidence that latent factors are necessarily caused by those variables through a new “causal measurement” mechanism.

When Arrows Tell Time

Unidirectional arrows might merit causal interpretation when the model's specification is tethered to information by the study design that distinguishes “before” from “after”, and the arrow conveys information about temporal direction. In a longitudinal model, for example, a path drawn from $\boldsymbol{X}$ at Time_1 to $\boldsymbol{Y}$ at Time_2 encodes a hypothesis that earlier variation in $\boldsymbol{X}$ helps predict later variation in $\boldsymbol{Y}$ over and above their autoregressive carry-over; because the variables are observed at different time points, the arrow corresponds with genuine temporal precedence, and its coefficient becomes a test of a putatively prospective effect. Similarly, in experimental or quasi-experimental designs, arrows emanating from a randomized treatment indicator (or a strong natural experiment) toward downstream outcomes inherit the study design's exogenous shock: here, causal inference rests on the randomly assigned manipulation, not on the algebra of the loading matrix. In both cases the graphical directionality is backed by study design features — temporal ordering, intervention, or an instrument that breaks reverse causation — so the arrow depicts more than a covariance constraint; it marks a falsifiable temporal or causal hypothesis whose validity can be probed through model fit, sensitivity checks, comparison to alternative models and, ideally, replication in independent samples.

Although ACE items refer to events that (by definition) occurred years before their hypothesized outcomes, their temporal precedence does not license a causal reading of the arrows that point from those items to a latent adversity factor in a formative specification. Retrospective ACE reports are themselves susceptible to current mood, personality, socially desirable responding, recall bias, demand characteristics, passive gene-environment correlation, and other unmeasured variables, so the observed covariation between ACE checklists and downstream outcomes can reflect those potential confounders — not an unambiguous unidirectional effect of childhood events accumulating adversity. The formative arrows are fitted to the same cross-sectional covariance matrix as any reflective arrows and, therefore, inherit its non-causal nature. Plus, treating the formative arrows as causal adds nothing beyond what is already established outside the model (historical ordering of events) while risking over-interpretation of coefficients.

The Reflective Model as Representational Measurement

Formative models arguably appeal to the desire for a measurement paradigm in which measurands or latent variables function as representations or summaries of observed variables. That is, conceptually what is sought is a mapping:

$$\boldsymbol{x} \Rightarrow \boldsymbol{f}$$

from the observed ***x*** variables to the latent factors ***f***, where the ***f*** represent, summarize, substitute or "stand in" for the ***x***. Here the arrow serves a mapping or "represent with" function from observed indicators ***x*** to the composite ***f***, not any causal influence. Related, "if a construct is composed of its measures, such that the measures are considered parts of the construct, then it makes little sense to treat the relationships between the construct and measures as causal…and if one variable is part of another, then their association is a type of part–whole correspondence, not a causal relationship between distinct entities"[6].

With ACEs, for example, we might seek a variable f that summarizes the phenomena x we see as constituting ACEs. As explained elsewhere[29], it is in fact reflective models, not formative models, that serve this representational function. The residual variances Ψ in a common factor model quantify the loss, in an information-theoretic sense, that one incurs by substituting ***f*** for ***x***. Mathematically, whenever one substitutes one variable for another, there must be some amount of information loss in doing so (unless the one variable is perfectly colinear with the other), and the reflective common factor model quantifies this loss. If, for example, ACE phenomena do not cohere well in the sense of fitting a reflective model, informationally, it suggests the ACE phenomena are better off modeled individually, with more factors, or with some other representational model.

That reflective common factor models serve a representational role in measurement further underscores how arrows in path diagrams can be misleading if interpreted causally. A better interpretation of the path diagram arrows, perhaps, is that one variable "accounts for" another (i.e., summarizes variance). Put differently, path diagram arrows in structural equation models, particularly at the cross-section, are model constraints, not automatic causal directives. Again, issues of causality involving the "real world" variables of interest are distinct from the models we use to represent them, including the variables from measurement models we use to represent the individuals and families we study.

A Causal Mirage on a Reflective Bedrock

In sum, in a cross-sectional study, treating the directional arrows in a path diagram as a suggestion, assumption, or poof that the latent factor causes variation in the items commits a conceptual error parallel to equating correlation with causation. Just as a nonzero correlation tells us only that two variables share variance, the loadings of the common factor merely partition the covariance matrix into a common part ($\Lambda\Phi\Lambda^T$) and a specific part (Ψ). Inferring causality from the arrow direction in a corresponding path diagram is no more defensible than claiming a causal pathway from ice cream sales to crime rates because they co-occur on hot days. Causal interpretation requires additional design features, not the sign convention of the loadings in a statistical model.i

The case for modeling adverse childhood experiences using a reflective model is decisive on both statistical and substantive grounds. First, model fit statistics cannot adjudicate whether ACE items "cause" a latent adversity construct — any formative block that reproduces the observed covariances will yield identical fit to a suitably specified reflective model, rendering model fit statistics moot. Second, securing identification for a purely formative specification requires additional constraints or auxiliary anchors, inflating analytic complexity and researcher degrees of freedom without delivering clear empirical insight. Third, the asymmetry in

[i] It may also be helpful to clarify the claim that "common factor models feature a local independence assumption in that any correlation between the indicators is assumed to be due to the causal influence of the underlying latent factor".[5] Local independence is a *modeling convenience*, not a causal dictum. It states only that—*conditional on the modeled factors*—residual correlations are fixed to zero unless the analyst frees them. Any remaining covariance can be accommodated by allowing correlated residuals, so the assumption neither demands nor proves the latent factor is a causal source of the indicators' association.

identification leans decisively toward reflection: a reflective factor persists when you drop "causal" indicator paths, whereas a formative construct vanishes without ad-hoc fixes or reflective anchors. Moreover, because ACEs naturally co-occur[21-23], reflective models harness this interdependence to produce stable scoring rules and directly quantify residual variance (Ψ), whereas formative weights vie to explain the same shared variance — amplifying standard errors. As shown in Appendix B, the correlations among ACE items from a highly cited study[21] are "meritorious" (.90 > *KMO* > .80) for a reflective factor analysis[30]. Finally, recognizing that estimated factor scores are themselves formative re-expressions of reflective parameters dissolves a practical rationale for a separate formative model: if all that is needed is an aggregated composite, the reflective framework already offers fit statistics, invariance testing, and robust scoring without the fragility and interpretive ambiguity[11] inherent to "causal" indicator specifications.

If policy makers, researchers, and clinicians require a single, portable metric of childhood adversity, cumulative stress[31], or cumulative advantage[32], the reflective factor model offers the clearest path. A formative construct whose definition drifts with each study's biomarker set or outcome battery risks re-introducing precisely the chaos that calls for standardization in the first place. Doing so does not deny that ACEs are causal (we believe they are); it simply keeps the ontological bookkeeping transparent. Researchers and other interested parties looking to measurements as representational summaries of a set of phenomena are better served by a reflective model anyway, theoretically and methodologically.

Although this paper adopts a skeptical view of formative indicator models, particularly "causal" indicator models, our purpose is not to dismiss or disparage the scholars who have advanced them. Rather, by iterating existing critiques and introducing additional considerations, we join a long line of researchers who have expounded objections to formative measurement[1,3,4,6,33,34]. We invite investigators weighing the use of formative models — particularly in the context of ACEs — to reflect on these objections, as we believe such scrutiny exposes the enticing yet ultimately false allure of "causal" indicator approaches.

Conflict of Interests

The authors have no conflict of interest to disclose.

Acknowledgements

We are grateful to Dr. Anthong Ong for providing valuable feedback on an earlier draft of this manuscript.

References

1. Howell RD, Breivik, E., & Wilcox, J. B. Reconsidering formative measurement. *Psychological Methods*. 2007;12(2):205–218.

2. Bollen KA. Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox. *Psychological Methods*. 2007;12(2):219–228.

3. Bagozzi RP. On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods*. 2007;12(2):229–237.

4. Howell RD, Breivik, E., & Wilcox, J. B. Is formative measurement really measurement? Reply to Bollen (2007) and Bagozzi (2007). *Psychological Methods*. 2007;12(2):238–245.

5. Moriarity DP, & Slavich, G. M. Theoretical and Quantitative Disconnect When Modeling Adverse Childhood Experiences Using a Common Factor Framework: An Argument for Causal Indicator Models in Stressor Research. *Child Development*. 2025:1-5.

6. Edwards JR. The fallacy of formative measurement. *Organizational Research Methods*. 2011;14(2):370-388.

7. Bedeian AG, Day, D. V., & Kelloway, E. K. Correcting for measurement error attenuation in structural equation models: Some important reminders. *Educational and Psychological Measurement*. 1997;57:785-799.

8. Blalock HM. *Causal inferences in nonexperimental research*. Chapel Hill: University of North Car-
olina Press; 1964.

9. Borsboom D, Mellenbergh, G. J., & Van Heerden, J. The theoretical status of latent variables. *Psychological Review*. 2003;110:203-219.

10. Diamantopoulos A, Riefler, P., & Roth, K. P. Advancing formative measurement models. *Journal of Business Research*. 2008;61:1203-1218.

11. Jarvis CB, MacKenzie, S. B., & Podsakoff, P. M. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*. 2003;30(2):199-218.

12. MacCallum R, & Browne, M. W. The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin*. 1993;114:533-541.

13. Podsakoff NP, Shen, W., & Podsakoff, P. M. The role of formative measurement models in strategic
management research: Review, critique and implications for future research. In: D. J. Ketchen DDB, ed. *Research methodology in strategy and management*. Elsevier.; 2006:197-252.

14. Bollen KA. *Structural Equations with Latent Variables*. John Wiley & Sons; 1989.

15. Thurstone LL. *The Vectors of Mind*. University of Chicago Press.; 1935:3–12.

16. Thomson GH. The definition and measurement of" g"(general intelligence). *Journal of Educational Psychology*. 1935;26(4):241.

17. McNeish D, Wolf MG. Thinking twice about sum scores. *Behavior research methods*. 2020:1-19.

18. Diamantopoulos A, & Winklhofer, H. M. Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research*. 2001;38:269-277.

19. Reilly T. A necessary and sufficient condition for identification of confirmatory factor analysis models of factoral complexity one. *Sociological Methods & Research*. 1995;23:421-441.

20. Davis WR. The FCI rule of identification for confirmatory factor analysis: A general sufficient condition. *Sociological Methods & Research*. 1993;21:403-437.

21. Dong M, Robert F. Anda, Vincent J. Felitti, Shanta R. Dube, David F. Williamson, Theodore J. Thompson, Clifton M. Loo, and Wayne H. Giles. The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Child Abuse & Neglect*. 2004;28(7):771-784.

22. Lian J, Kim M. Kiely, and Kaarin J. Anstey. Cumulative risk, factor analysis, and latent class analysis of childhood adversity data in a nationally representative sample. *Child Abuse & Neglect*. 2022;125:105486.

23. Yu Z, Tejashwini Alalamath, Sara Johnson, Yiying Xiong, and Lin Wang. Network analysis of adverse and positive childhood experiences in a national sample of university students. *Social Science & Medicine*. 2025:17792.

24. Bollen KA, & Davis, W. R. Causal indicator models: Identification, estimation, and testing. *Structural Equation Modeling: A Multidisciplinary Journal*. 2009;16(3):498-522.

25. Costner HL. Theory, deduction, and the rules of correspondence. *American Journal of Sociology*. 1969;75:245-263.

26. MacKenzie SB, Podsakoff, P. M., & Jarvis, C. B. The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *Journal of applied psychology*. 2005;90(4):710-730.

27. Heise DR. Employing nominal variables, induced variables, and block variables in path analysis. *Sociological Methods & Research*. 1972;1:147-173.

28. Aguirre-Urreta MI, Rönkkö, M., & Marakas, G. M. Omission of causal indicators: Consequences and implications for measurement. Measurement: Interdisciplinary Research and Perspectives. *Measurement: Interdisciplinary Research and Perspectives*. 2016;14(3):75-97.

29. Markon KE. Reliability as Lindley information. *Multivariate Behavioral Research*. 2023;58(4):815-842.

30. Kaiser HF. An index of factorial simplicity. *Psychometrika*. 1974;39(1):31-36.

31. Mann FD, Cuevas AG, Krueger RF. Cumulative stress: A general "s" factor in the structure of stress. *Social Science & Medicine*. 2021;289:114405, 1-10.

32. Ong AD, & Mann, F. D. Cumulative social advantage across the lifespan: Examining the convergent and predictive validity of a multidimensional hierarchical construct for health and longevity. *American Psychologist*. 2025;

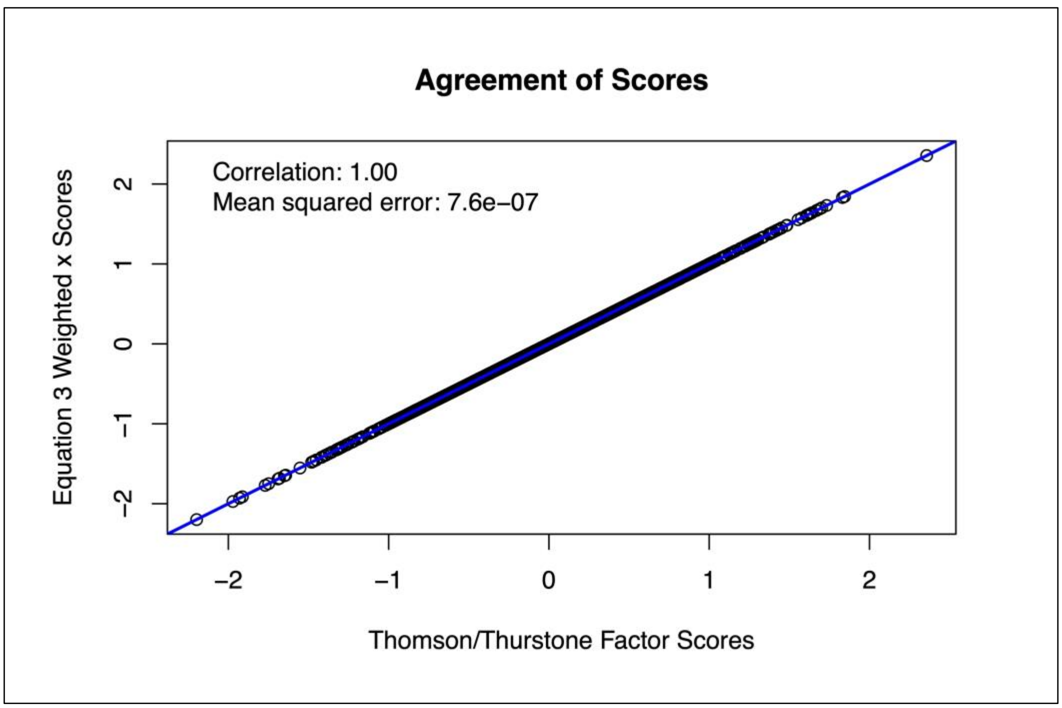
33. Iacobucci D. Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*. 2010;20:90-98.

34. Wilcox JB, Howell, R. D., & Breivik, E. Questions about formative measurement. *Journal of Business Research*. 2008;61:1219-1228.

Appendix A

```
# Simulate a one-factor model dataset with 4 indicators
set.seed(0.57721)        # for reproducibility
n <- 1000                # number of observations

# Specify true loadings for 4 indicators
lambda <- c(0.8, 0.7, 0.6, 0.5)

# Simulate latent factor f ~ N(0,1)
f <- rnorm(n, mean = 0, sd = 1)

# Simulate unique errors so that Var(x_i) = 1
e1 <- rnorm(n, sd = sqrt(1 - lambda[1]^2))
e2 <- rnorm(n, sd = sqrt(1 - lambda[2]^2))
e3 <- rnorm(n, sd = sqrt(1 - lambda[3]^2))
e4 <- rnorm(n, sd = sqrt(1 - lambda[4]^2))

# Create indicators
x1 <- lambda[1] * f + e1
x2 <- lambda[2] * f + e2
x3 <- lambda[3] * f + e3
x4 <- lambda[4] * f + e4

# Combine indicators into a data frame
df <- data.frame(x1, x2, x3, x4)

# Fit a one-factor reflective CFA with lavaan
# Run install.packages if lavaan package is not already installed
# install.packages('lavaan')
library(lavaan)

mod <- '
  f =~ x1 + x2 + x3 + x4
'
fit <- cfa(mod, data = df)

# Examine fit statistics (out of habit)
key_fits <- fitMeasures(fit,
                        c("chisq","df","pvalue","cfi","tli","rmsea",
                          "rmsea.ci.lower","rmsea.ci.upper","srmr"))
key_fits

# Extract parameter matrices
Lambda <- inspect(fit, "est")$lambda   # 4 × 1 loadings
Theta  <- inspect(fit, "est")$theta    # 4 × 4 residual-variance
Phi    <- inspect(fit, "est")$psi      # 1 × 1 latent-variance

# Compute the regression-weight row vector
# Gamma_t is 1 x 4:
Gamma_t <- solve(
  t(Lambda) %*% solve(Theta) %*% Lambda + solve(Phi)
) %*% t(Lambda) %*% solve(Theta)
```

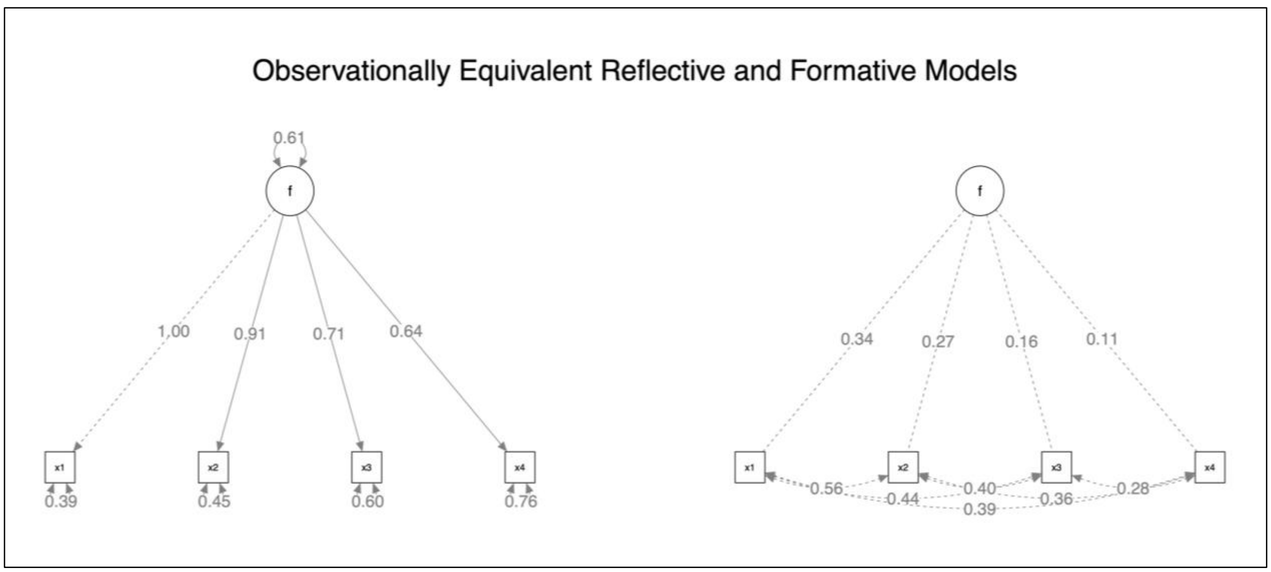
```
# Ensure a 4 × 1 column vector
gamma_vec <- as.numeric(t(Gamma_t))    # length-4 vector

# Compute factor scores via the linear predictor
X <- as.matrix(df[, c("x1", "x2", "x3", "x4")])  # 1000 × 4 data matrix
f_manual <- X %*% gamma_vec                        # 1000 × 1 predicted score

# Save lavaan's built-in regression-based factor scores
f_lavaan <- lavPredict(fit)                        # 1000 × 1 vector

# Quantify agreement between "hand" calculated vs. default/automated scores
correlation <- cor(f_manual, f_lavaan)
mse         <- mean((f_manual - f_lavaan)^2)

# Create Scatter Plot of Score Agreement
plot(f_lavaan, f_manual,
     xlab = "Thomson/Thurstone Factor Scores",
     ylab = "Equation 3 Weighted x Scores",
     main  = "Agreement of Scores")
abline(0, 1, col = "blue", lwd = 2)

legend("topleft",
       legend = c(
         sprintf("Correlation: %.2f", correlation),
         sprintf("Mean squared error: %.2g", mse)
       ),
       bty = "n")
```




```
# Perfect correlation and approximately zero mean squared error demonstrates
equation (4)
```

```
# Save implied covariance matrix from reflective model
Sigma_hat <- fitted(fit)$cov

# Solve Var(zeta) so that Var(f) = Gamma'SigmaGamma + Var(zeta)
zeta_var  <- Phi - as.numeric(t(gamma_vec) %*% Sigma_hat %*% gamma_vec)

# Re-fit models to Sigma_hat
# (a) Reflective
fit_ref <- cfa(mod,
                sample.cov  = Sigma_hat,
                sample.nobs = n)

# (b) Formative
form_mod  <- sprintf(
  "f <~ %g*x1 + %g*x2 + %g*x3 + %g*x4\nf ~~ %g*f",
  gamma_vec[1], gamma_vec[2], gamma_vec[3], gamma_vec[4], zeta_var
)

fit_form <- sem(form_mod,
                sample.cov  = Sigma_hat,
                sample.nobs = n)

# Confirm identical implied covariance via the difference between the
# matrices because fit statistics are not available for the formative

max_diff <- max(abs(fitted(fit_ref)$cov - fitted(fit_form)$cov))
cat("Max absolute difference in implied variance-covariance matrix:",
max_diff, "\n")

# Max absolute difference in implied variance-covariance matrix: 1.474721e-08
# The difference is essentially zero (to machine-level precision)

# Create path diagrams of the reflective & formative models
library(semPlot)

# Set up 1×2 panels, inner margins for each plot, and extra outer margin
par(
  mfrow = c(1, 2),        # 1 row, 2 columns
  mar   = c(0, 1, 0, 1),  # bottom, left, top, right (inner margins)
  oma   = c(0, 0, 3, 0)   # outer margin: bottom, left, top, right
)

# First panel: reflective model
semPaths(
  fit_ref,
  whatLabels     = "est",
  layout         = "tree",
  style          = "ram",
  shapeMan       = "rectangle",
  shapeLat       = "circle",
  exoCov         = TRUE,
  intercepts     = FALSE,
  residuals      = TRUE,
```

```
  edge.label.cex = 1.25
)

# Second panel: formative model
semPaths(
  fit_form,
  what           = "paths",
  whatLabels     = "est",
  layout         = "tree",
  style          = "ram",
  shapeMan       = "rectangle",
  shapeLat       = "circle",
  exoCov         = TRUE,
  intercepts     = FALSE,
  residuals      = FALSE,
  edge.label.cex = 1.25
)

# Figure title
mtext(
  "Observationally Equivalent Reflective and Formative Models",
  outer = TRUE,      # use the outer margin
  cex   = 1,       # title size
  line  = 1          # how far down from the top edge
)
```



Appendix B

```
# The following code creates a matrix of odds ratios (ORs) that report
# relations among ten ACEs from a highly cited study. ORs are then
# transformed to obtain Cohen’s d, which are then converted to
# Pearson’s r, and the suitability of the data for a reflective factor
# analysis is evaluated with Bartlett’s test and the Kaiser-Meyer-Olkin
# (KMO) index.

# Source of ORs: Dong, M., Anda, R. F., Felitti, V. J., Dube, S. R.,
# Williamson, D. F., Thompson, T. J., ... & Giles, W. H. (2004). The
# interrelatedness of multiple forms of childhood abuse, neglect, and
# household dysfunction. Child Abuse & Neglect, 28(7), 771-784.

# ACE labels
ace_names <- c(
  "em_abuse","ph_abuse","sex_abuse",
  "em_neglect","ph_neglect",
  "par_sep","sub_abuse","mh_house","dom_viol","crime"
)

# Marginal prevalences (Table 1)
p_i <- c(
  em_abuse   = 0.102,
  ph_abuse   = 0.264,
  sex_abuse  = 0.210,
  em_neglect = 0.148,
  ph_neglect = 0.099,
  par_sep    = 0.130,
  sub_abuse  = 0.282,
  mh_house   = 0.203,
  dom_viol   = 0.241,
  crime      = 0.060
)

# Initialize OR_mat
ace_names <- c(
  "em_abuse","ph_abuse","sex_abuse",
  "em_neglect","ph_neglect",
  "par_sep","sub_abuse","mh_house","dom_viol","crime"
)
OR_mat <- matrix(NA_real_, 10, 10,
                 dimnames = list(ace_names, ace_names))
diag(OR_mat) <- 1

# Helper function to fill [i,j] and [j,i]
fill_or <- function(i, j, val) {
  OR_mat[i, j] <<- val
  OR_mat[j, i] <<- val
}

# Get ORs from Table 2
fill_or("ph_abuse",   "em_abuse",    17.7)
fill_or("sex_abuse",  "em_abuse",     3.0)
fill_or("ph_neglect", "em_abuse",     6.3)
```

```
fill_or("em_neglect", "em_abuse",    12.8)

fill_or("sex_abuse",  "ph_abuse",     2.4)
fill_or("em_neglect", "ph_abuse",     5.1)
fill_or("ph_neglect", "ph_abuse",     3.7)

fill_or("em_neglect", "sex_abuse",    2.5)
fill_or("ph_neglect", "sex_abuse",    2.5)

fill_or("ph_neglect", "em_neglect",  12.2)

fill_or("par_sep",    "em_abuse",     2.0)
fill_or("sub_abuse",  "em_abuse",     2.9)
fill_or("mh_house",   "em_abuse",     4.2)
fill_or("dom_viol",   "em_abuse",     5.9)
fill_or("crime",      "em_abuse",     2.7)

fill_or("par_sep",    "ph_abuse",     2.2)
fill_or("sub_abuse",  "ph_abuse",     2.1)
fill_or("mh_house",   "ph_abuse",     2.8)
fill_or("dom_viol",   "ph_abuse",     4.7)
fill_or("crime",      "ph_abuse",     2.5)

fill_or("par_sep",    "sex_abuse",    2.0)
fill_or("sub_abuse",  "sex_abuse",    2.0)
fill_or("mh_house",   "sex_abuse",    2.1)
fill_or("dom_viol",   "sex_abuse",    2.5)
fill_or("crime",      "sex_abuse",    2.4)

fill_or("par_sep",    "em_neglect",   2.7)
fill_or("sub_abuse",  "em_neglect",   2.5)
fill_or("mh_house",   "em_neglect",   3.3)
fill_or("dom_viol",   "em_neglect",   4.0)
fill_or("crime",      "em_neglect",   2.2)

fill_or("par_sep",    "ph_neglect",   2.6)
fill_or("sub_abuse",  "ph_neglect",   3.0)
fill_or("mh_house",   "ph_neglect",   3.4)
fill_or("dom_viol",   "ph_neglect",   4.6)
fill_or("crime",      "ph_neglect",   2.5)

fill_or("sub_abuse",  "par_sep",      2.9)
fill_or("mh_house",   "par_sep",      2.5)
fill_or("dom_viol",   "par_sep",      3.9)
fill_or("crime",      "par_sep",      2.6)

fill_or("mh_house",   "sub_abuse",    2.7)
fill_or("dom_viol",   "sub_abuse",    5.9)
fill_or("crime",      "sub_abuse",    3.3)

fill_or("dom_viol",   "mh_house",     2.8)
fill_or("crime",      "mh_house",     3.3)

fill_or("crime",      "dom_viol",     3.2)
```

```
# Sanity-check
if(any(is.na(OR_mat))) {
  miss <- which(is.na(OR_mat), arr.ind=TRUE)
  stop("Missing OR for pairs:\n",
       paste0(rownames(OR_mat)[miss[,1]], "-",
              colnames(OR_mat)[miss[,2]], collapse="\n"))
}

# View the fully populated OR_mat
print(OR_mat)

# Sanity check: diagonal should be 1
diag(OR_mat) == 1

# Now convert OR-->Cohen's d-->Pearson's r
d_mat  <- log(OR_mat) * sqrt(3)/pi
rho_or <- d_mat / sqrt(d_mat^2 + 4)
diag(rho_or) <- 1
rho_or

# can force PD (not necessary)
library(Matrix)
rho_pd <- as.matrix(nearPD(rho_or)$mat)

# Sample size
n = 8629

# Codes for Bartlett's test and KMO obtained from:
# https://data-mining-tutorials.blogspot.com/2013/01/pca-using-r-kmo-# index-
and-bartletts-test.html

# Bartlett's test
p    <- ncol(rho_pd)
chi2 <- -(n - 1 - (2*p + 5)/6) * log(det(rho_pd))
df   <- p*(p - 1)/2
pval <- pchisq(chi2, df, lower.tail = FALSE)
cat("Bartlett's test =", chi2, " df =", df, " p =", pval, "\n")

#Bartlett's test = 22759.71  df = 45  p = 0

# KMO
invR     <- solve(rho_pd)
A        <- diag(p)
for(i in 1:(p-1)) for(j in (i+1):p) {
  A[i,j] <- A[j,i] <- -invR[i,j] / sqrt(invR[i,i] * invR[j,j])
}
R2       <- rho_pd^2
partial2 <- A^2
kmo_overall <- sum(R2[upper.tri(R2)]) /
  (sum(R2[upper.tri(R2)]) + sum(partial2[upper.tri(partial2)]))
kmo_by_var <- sapply(1:p, function(i) {
  sum(R2[i, -i]) / (sum(R2[i, -i]) + sum(partial2[i, -i]))
})
names(kmo_by_var) <- colnames(rho_pd)
```

```
cat("Overall KMO:", round(kmo_overall,3), "\n")
print(round(kmo_by_var,3))

# Overall KMO: 0.851
#  em_abuse   ph_abuse  sex_abuse em_neglect ph_neglect    par_sep  sub_abuse
#     0.800      0.824      0.936      0.828      0.864      0.857      0.852
#  mh_house   dom_viol      crime
#     0.908      0.873      0.884

# Now, let's confirm suitability using tetrachoric correlations
library(psych)
tetra_mat <- matrix(NA_real_,
                    nrow = nrow(OR_mat),
                    ncol = ncol(OR_mat),
                    dimnames = dimnames(OR_mat))

for(i in seq_len(nrow(OR_mat))) {
  for(j in seq_len(i)) {
    Pi <- p_i[i]; Pj <- p_i[j]; OR <- OR_mat[i,j]
    a <-  OR - 1
    b <- -(OR*(Pi+Pj) + 1 - Pi - Pj)
    c <-  OR * Pi * Pj

    roots      <- polyroot(c(c, b, a))
    real_roots <- Re(roots)[abs(Im(roots))<1e-8]
    lower      <- max(0, Pi+Pj-1)
    upper      <- min(Pi, Pj)
    p11        <- real_roots[real_roots>=lower & real_roots<=upper][1]

    p10 <- Pi - p11; p01 <- Pj - p11; p00 <- 1 - Pi - Pj + p11

    tab <- matrix(round(c(p11, p10, p01, p00)*n),
                  nrow = 2, byrow = TRUE,
                  dimnames = list(c("1","0"), c("1","0")))

    # Extract the scalar tetrachoric from the list
    tetra_r <- as.numeric(tetrachoric(tab, correct = FALSE)$rho)

    tetra_mat[i, j] <- tetra_r
    tetra_mat[j, i] <- tetra_r
  }
}

diag(tetra_mat) <- 1
tetra_mat

# Bartlett's test
p    <- ncol(tetra_mat)
chi2 <- -(n - 1 - (2*p + 5)/6) * log(det(tetra_mat))
df   <- p*(p - 1)/2
pval <- pchisq(chi2, df, lower.tail = FALSE)
cat("Bartlett's test =", chi2, " df =", df, " p =", pval, "\n")

# Bartlett's test = 34295.43  df = 45  p = 0
```

```
# KMO
invR     <- solve(tetra_mat)
A        <- diag(p)
for(i in 1:(p-1)) for(j in (i+1):p) {
  A[i,j] <- A[j,i] <- -invR[i,j] / sqrt(invR[i,i] * invR[j,j])
}
R2       <- tetra_mat^2
partial2 <- A^2
kmo_overall <- sum(R2[upper.tri(R2)]) /
  (sum(R2[upper.tri(R2)]) + sum(partial2[upper.tri(partial2)]))
kmo_by_var <- sapply(1:p, function(i) {
  sum(R2[i, -i]) / (sum(R2[i, -i]) + sum(partial2[i, -i]))
})
names(kmo_by_var) <- colnames(tetra_mat)

cat("Overall KMO:", round(kmo_overall,3), "\n")
print(round(kmo_by_var,3))

# Overall KMO: 0.833
# print(round(kmo_by_var,3))
#  em_abuse   ph_abuse  sex_abuse em_neglect ph_neglect    par_sep  sub_abuse
#     0.775      0.813      0.957      0.795      0.845      0.820      0.827
#  mh_house   dom_viol      crime
#     0.916      0.852      0.899

# To formally assess the factor structure of ACE items, the raw data
# should be obtained and analyzed. Nevertheless, the correlations among
# the ACE items are sufficient to be considered "meritorious" for a
# reflective factor analysis
```