



How to Make Recommendations for Educational Practice from Correlational Data Using Structural Equation Models

Denis Dumas¹ · Peter Edelsbrunner²

Accepted: 8 April 2023 / Published online: 19 April 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

In this commentary, we outline a five-phase process by which recommendations for educational practice can be distilled from correlational data using structural equation modeling (SEM). First, meta-theoretical beliefs associated with latent variables—that mental attributes cause behavior and can therefore be measured indirectly by observing multiple indicators of that behavior—must be adopted and made explicit. Next, an SEM must be formulated with relevant pathways and covariates that exhaustively represent our theoretical knowledge and assumptions about the structure of the psychological phenomena being studied. Third, model-data-fit indices and estimated parameters associated with the SEM should be carefully interpreted. Fourth, the model should be replicated across educational contexts, and any necessary changes should be incorporated into the relevant psychological theory. Fifth, the results of multiple studies can then be interpreted together with other sources of evidence as a basis for communicating our current theoretical understanding and caveats to practitioners. We also point out that educational recommendations should likely never be entirely prescriptive, and instead lie on a continuum of specificity based on the strength of the evidence.

Keywords Structural equation modeling · Recommendations for practice · Research design · Causation · Correlational Data

This paper is a commentary on Brady and colleagues' (2023) "How scientific is educational psychology research? The increasing trend of squeezing causality and recommendations from non-intervention studies".

✉ Denis Dumas
denis.dumas@uga.edu

¹ Department of Educational Psychology, University of Georgia, 624 Aderhold Hall, 110 Carlton St, Athens, GA 30602, USA

² Department of Humanities, Social, and Political Sciences, ETH Zürich, Zurich, Switzerland

At times, educational psychologists step into the role of research translators, in which we attempt to distill reasoned and reasonable recommendations for educational practice or policy from our scientific work (c.f., Alexander, 2013; Mayer, 2003; Renkl, 2013). For example, educational psychologists have recently forwarded practical recommendations concerning testing and measurement issues (e.g., McNeish & Dumas, 2019); instructional techniques (e.g., Reeve & Cheon, 2021); and diversity, equity, and inclusion efforts (e.g., Juvonen et al., 2019), among many areas. However, some in the educational psychology literature (e.g., Robinson et al., 2007, 2013) have argued that we may be inadvertently steering practitioners wrong by overstepping the data available to us.

A strongly worded paper in this line came from Brady and colleagues (2023). These scholars contended that educational psychologists should solely infer recommendations for practice from data drawn from experimental intervention studies. Although we agree with Brady and colleagues' most fundamental point: that practical recommendations should be forwarded in a responsible way based on the evidence at hand; we do not agree that intervention research is the only form of inquiry that provides the evidence needed for practical recommendations. Instead—following in-line with other areas of science such as tobacco smoking (Sasco et al., 2004) or the atmospheric greenhouse effect (Nissani, 1996)—we argue that it is possible for researchers to make carefully considered recommendations based on multiple sources of observational and correlational data.

In the context of educational psychology, we suggest that structural equation modeling (SEM; Bentler, 1980; Hoyle, 2023; Jöreskog, 1978) is a methodological paradigm that can support justified recommendations for practice based on correlational data. We outline a five-phase process in which SEM, under consideration of recent methodological developments in the modeling of correlational data, can be used to carefully derive implications for practice and test their generalizability. This way, as the field moves forward, educational psychologists can continue to fulfill their roles as responsible research translators, while also taking full advantage of the psychological meaning in their correlational data.

Phase One: Adopt Meta-Theoretical Beliefs About the Mind

A key meta-theoretical belief among psychologists who utilize SEM is that mental attributes can be indirectly measured by observing behavior (Bollen, 2002; Bollen & Hoyle, 2023; Borsboom, 2008). This meta-theoretical understanding can be tracked back more than a century, at least to the work of Charles Spearman (e.g., 1907) who posited that variance in observed behavior was underlain by two sources: latent psychological attributes that cause the behavior, and random measurement error inherent to the process of doing psychological research. This deceptively simple meta-theoretical belief can still be observed in SEM today, where behavioral indicators are endogenous to (i.e., affected by) their hypothesized latent psychological causes, and to their corresponding error terms (Hoyle, 2023; Mueller, & Hancock, 2019). To put it simply, SEM researchers believe that the mind causes behavior, and therefore behavior can be used to make inferences about the mind.

In addition, SEM meta-theoretically posits that the variance and covariance among behavioral indicators can be mathematically reduced to a smaller number of vectors or dimensions that represent the latent psychological variables being modeled (Thurstone, 1940). All leftover variance that is not represented in this smaller number of psychological dimensions is assumed to be caused by measurement error. Notably, there are other meta-theoretical beliefs about psychological phenomena; for example, some latent constructs could be caused or formed by indicator variables instead of the other way round (Schuberth, 2021), or behaviors captured by indicators could affect each other directly rather than indicate a common latent variable (van der Maas et al., 2006). In order to adopt an SEM-based research agenda, the first step is to decide which of these meta-theoretical assumptions best represents the psychological phenomena of interest. Although we limit the present discussion to working with latent variables that causally explain observed behavior, constructs for which other kinds of meta-theories are more appropriate can also be integrated into SEM (Epskamp et al., 2017; Schuberth, 2021).

Phase Two: Configure a Theoretically Meaningful Model

Next, a model must be built that, as faithfully as possible, represents how the phenomenon of interest unfolds. In this theoretical phase, a research team will closely interrogate the existing literature to form hypotheses about the interrelations among the latent variables in their model. What patterns of prediction and causes (Kline, 2023), what moderation (Kelava & Brandt, 2023), and what mediation may be occurring (Gonzalez et al., 2023)? What are the important covariates that also exist in this space—and help to cause variance and covariance in the behavioral indicators—that must be included in the model, and the inclusion of what other covariates might be misleading (Pearl, 2023)? If the possibility of causal conclusions is in focus, then longitudinal data that may be particularly helpful for integrating model structures into SEM that ensure biased results are less likely (Lüdtke & Robitzsch, 2022; McNeish et al., 2022). All pathways in a hypothesized model should have past citations or reasonably founded theoretical justifications associated with them, and ways in which your model deviates from the existing literature should be made explicit during publication.

In this way, the model is configured to match the most rigorous extant understanding of the phenomena being studied. In cases where multiple model configurations are roughly equally supported by the existing literature, both can and should be tested to determine which fits the data better (Preacher and Yaremych, 2023). To put it another way, the configuration of the SEM is meant to be an as-direct-as-possible translation of psychological theory into a statistical form, which is an important demand for making any statistical model informative (Robinaugh et al., 2021). In cases where researchers are not able to represent existing theory fully and faithfully in their model, perhaps because they did not measure all relevant psychological variables and covariates, the ways in which the model misses these theoretically relevant components should be made explicit as limitations. The more relevant covariates that are available to include in the model, the more alternative explanations for

correlations in our data can be ruled out and indirect evidence for causation can be gathered (Reiss, 2015).

Phase Three: Interpret Model-Data-Fit Indices and Model Coefficients

For SEM researchers, it is the model's job to reproduce the patterns in the observed dataset by estimating values for all model parameters that best imply the observed variances and covariances. Fit indices are then computed and used to examine how well the model-estimated parameters manage to describe the variances and covariances of the observed variables (West et al., 2023). These fit indices can be used to evaluate whether the theoretical considerations that have been used to define an SEM have led to parameters that appropriately manage to describe the observed data.

Of note, which fit indices and cut-offs are most appropriate depends on the research context, for example on the nature and number of the variables in the model, and the amount of residual variance in the observed variables (Hancock & Mueller, 2011; Heene et al., 2011). Vast literature is available to guide researchers in selecting from available fit indices and to determine which cut-offs on these indices might indicate mild or substantial deviations of the empirical data-patterns from the patterns implied by the model (see e.g., Greiff & Heene, 2017; McNeish & Wolf, 2021; West et al., 2023).

If a model fits the data well, estimated coefficients in that model can hold rich psychological information about students. For instance, through the model loadings, SEMs can depict how closely behavioral indicators relate to their corresponding latent variables (Bollen & Hoyle, 2023). Over and above the actual configuration of the structural model, which describes how the latent variables are hypothesized to interrelate, the structural coefficients capture the direction and degree of those interrelations. If a researcher has adopted the meta-theoretical beliefs described above and has configured their model to as-closely-as-possible represent psychological theory, these structural coefficients represent the essential make-up of the human mind vis-à-vis the theory being tested.

Phase Four: Replicate Across Contexts

In educational psychology, phenomena typically vary across contexts (Berliner, 2002; Hedges, 2013; Plucker & Makel, 2021). For this reason, even the most well-supported models remain context-dependent theories and not laws in perpetuity. Appropriate fit across contexts supports a model's capability to be a reliable and valid tool for prediction and for deriving implications for practice. In contrast, contexts from which the available data do not fit the model can be taken as potential boundary conditions to the validity of the existing model. The model would then need to be changed (e.g., Harackiewicz et al., 2002) or extended (e.g., Merk et al., 2018; Wolff et al., 2019) to accommodate the new data.

In cases where the theory (and the SEM that represents it) needs to be amended based on newly available data from an additional educational context, the hypothesized reasons why those amendments were needed should also be incorporated into the theory. This opens the door to the possibility of testing the new wider theory by modeling patterns of variance and covariance across contexts via approaches such as multi-level (Heck & Reid, 2023), multi-group (Widaman & Olivera-Aguilar, 2023), meta-analytic (Cheung, 2023), local (Hildebrandt et al., 2016), or moderated SEM (Molenaar, 2021). This process of broadening what is known about the phenomena being studied, whether by identifying contexts in which the model holds as it is, or contexts where it needs to be amended, is the expansion of our scientific understanding in educational psychology. Some examples of this process within our field might be the Big-Fish-Little-Pond effect (Werts & Watley, 1969) or the g-factor of cognitive ability (Spearman, 1927), which have been replicated in many populations around the world but have failed to hold up in others (Guilford, 1964; Seaton et al., 2009; Warne & Burningham, 2019).

Importantly, SEM, as any other methodological approach, is not a technique in isolation. The findings from SEM-based studies should always be compared with evidence from other approaches to see whether under methodological pluralism, findings hold up and allow arriving at robust conclusions (Oreskes, 2019). When studies that have undertaken approaches such as controlled trials, inter- and intraindividual perspectives (which can both be integrated in SEM; see Asparouhov et al., 2018), and qualitative/mixed methods approaches all yield comparable results, then we can be sure that we have modeled robust phenomena rather than tendencies of specific approaches to yield specific patterns (Eid et al., 2023). Generalization across populations and across methods are both required to arrive at the most robust conclusions.

Phase Five: Support Practitioners in Reasoning About the Evidence

Whenever educational psychologists communicate with practitioners or policymakers, care and caution are warranted so as not to overstate the evidence. For instance, good recommendations should only involve the variables included in an SEM and not make any conjecture about unmeasured variables external to the model. In addition, the actual strength of the coefficients in the model should be carefully communicated so that practitioners can understand what the numbers mean in terms of educational practices and outcomes. The degree to which the model fits the data—if the fit is excellent or if it approaches the borderline of quality standards—should also be communicated, along with the extant contexts across which the model has been replicated. In this way, practitioners can receive information that allows them to reason about the strength of the evidence and make their own decision about whether to believe a theory and adopt it in their practice.

It is important to consider the educational background of the practitioners to whom we communicate the most frequently. For example, school psychologists, educational specialists, and classroom teachers, might have very different educational backgrounds and knowledge about the methods utilized in a study. Literature

on communicating statistical information (e.g., Schmidt et al., *in press*) and examples from Clearinghouses (WWC, 2022) can help in designing the communication of results from SEM-based research such that practitioners can validly interpret the information and can use it for evidence-informed decisions in their educational practice (Greisel et al., 2023).

Conclusion

Instead of completely avoiding the derivation of practical recommendations from observational or correlational data, we should do so in order not to lose potentially useful and valid information (Grosz et al., 2020), and we should bring the utmost care to this endeavor. Intervention studies and correlational designs have their own unique strengths and weaknesses when it comes to the validity of inferences. Eventually, whether practical implications can be drawn is not a yes/no question. All kinds of studies should be seen as lying on a continuum from “only observational statements possible” to “valid evidence for deducing recommendations for practice”. The better observational studies are designed, and their data analyzed in light of theory, taking into account appropriate covariates and harnessing modern modeling approaches, the more we can be sure to be further toward the latter end of this continuum.

References

- Alexander, P. A. (2013). In praise of (reasoned and reasonable) speculation: A response to Robinson et al.'s moratorium on recommendations for practice. *Educational Psychology Review*, 25(3), 303–308. <https://doi.org/10.1007/s10648-013-9234-2>
- Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 359–388.
- Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, 40(1), 8–34. <https://doi.org/10.1007/s11747-011-0278-x>
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31(1), 419–456.
- Berliner, D. C. (2002). Comment: Educational Research: The Hardest Science of All. *Educational Researcher*, 31(8), 18–20. <https://doi.org/10.3102/0013189X031008018>
- Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, 53(1), 605–634. <https://doi.org/10.1146/annurev.psych.53.100901.135239>
- Bollen, K. A., & Hoyle, R. H. (2023). Latent Variables in Structural Equation Modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 97–109. Guilford.
- Bollen, K. A., & Long, J. S. (1993). *Testing Structural Equation Models*. SAGE.
- Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research and Perspectives*, 6(1–2), 25–53. <https://doi.org/10.1080/15366360802035497>
- Brady, A. C., Griffin, M. M., Lewis, A. R., Fong, C. J., & Robinson, D. H. (2023). How scientific is educational psychology research? The increasing trend of squeezing causality and recommendations from non-intervention studies. *Educational Psychology Review*, 35(1), 37. <https://doi.org/10.1007/s10648-023-09759-9>
- Cheung, M. W.L. (2023). Structural Equation Modeling-Based Meta-Analysis. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 664–680. Guilford.

- Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research*, 27(1), 131–157. https://doi.org/10.1207/s15327906mbr2701_8
- Eid, M., Koch, T., & Geiser, C. (2023). Multitrait-multimethod models. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 349–266. Guilford.
- Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 82(4), 904–927. <https://doi.org/10.1007/s11336-017-9557-x>
- Feng, Y., & Hancock, G. R. (2022). Model-based incremental validity. *Psychological Methods*, 27, 1039–1060. <https://doi.org/10.1037/met0000342>
- Gibson, W. A. (1962). Class assignment in the latent profile model. *Journal of Applied Psychology*, 46, 399–400. <https://doi.org/10.1037/h0043541>
- Gonzalez, O., Valente, M. J., Cheong, J., & MacKinnon, D. P. (2023). Mediation/indirect effects in structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 409–426. Guilford.
- Greiff, S., & Heene, M. (2017). Why psychological assessment needs to start worrying about model fit. *European Journal of Psychological Assessment*, 33(5), 313–317. <https://doi.org/10.1027/1015-5759/a000450>
- Greisel, M., Wekerle, C., Wilkes, T., Stark, R., & Kollar, I. (2023). Pre-service teachers' evidence-informed reasoning: Do attitudes, subjective norms, and self-efficacy facilitate the use of scientific theories to analyze teaching problems? *Psychology Learning & Teaching*, 22(1), 20–38. <https://doi.org/10.1177/14757257221113942>
- Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science*, 15(5), 1243–1255.
- Guilford, J. P. (1964). Zero correlations among tests of intellectual abilities. *Psychological Bulletin*, 61(6), 401.
- Hancock, G. R., & Mueller, R. O. (2011). The reliability paradox in assessing structural relations within covariance structure models. *Educational and Psychological Measurement*, 71(2), 306–324. <https://doi.org/10.1177/0013164410384856>
- Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology*, 94, 638–645. <https://doi.org/10.1037/0022-0663.94.3.638>
- Heck, R. H., & Reid, T. (2023). Multilevel structural equation modeling: An overview. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 481–500. Guilford.
- Hedges, L. V. (2013). Recommendations for practice: Justifying claims of generalizability. *Educational Psychology Review*, 25(3), 331–337. <https://doi.org/10.1007/s10648-013-9239-x>
- Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, 16(3), 319–336. <https://doi.org/10.1037/a0024917>
- Hildebrandt, A., Lüdtke, O., Robitzsch, A., Sommer, C., & Wilhelm, O. (2016). Exploring factor model parameters across continuous variables with local structural equation models. *Multivariate Behavioral Research*, 51(2–3), 257–258.
- Hoyle, R. H. (2023). Structural equation modeling: An overview. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 3–16. Guilford.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183–202. <https://doi.org/10.1007/BF02289343>
- Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, 43(4), 443–477. <https://doi.org/10.1007/BF02293808>
- Juvenon, J., Lessard, L. M., Rastogi, R., Schacter, H. L., & Smith, D. S. (2019). Promoting social inclusion in educational settings: Challenges and opportunities. *Educational Psychologist*, 54(4), 250–270. <https://doi.org/10.1080/00461520.2019.1655645>
- Kelava, A., & Brandt, H. (2023). Latent interaction effects. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed). 427–446. Guilford.
- Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* (4th Ed). Guilford.
- Kline, R. B. (2023). Assumptions in structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 128–144. Guilford.

- Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling: A Multidisciplinary Journal*, 29(6), 888–907.
- Mayer, R. E. (2003). Learning environments: The case for evidence-based practice and issue-driven research. *Educational Psychology Review*, 15(4), 359–366. <https://doi.org/10.1023/A:1026179332694>
- McNeish, D., & Dumas, D. G. (2019). Scoring repeated standardized tests to estimate capacity, not just current ability. *Policy Insights from the Behavioral and Brain Sciences*, 6(2), 218–224. <https://doi.org/10.1177/2372732219862578>
- McNeish, D., & Wolf, M. G. (2021). Dynamic fit index cutoffs for confirmatory factor analysis models. *Psychological Methods*. <https://doi.org/10.1037/met0000425>
- McNeish, D., Harring, J. R., & Dumas, D. (2022). A multilevel structured latent curve model for disaggregating student and school contributions to learning. *Statistical Methods & Applications*. <https://doi.org/10.1007/s10260-022-00667-w>
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55(1), 107–122. <https://doi.org/10.1007/BF02294746>
- Merk, S., Rosman, T., Muis, K. R., Kelava, A., & Bohl, T. (2018). Topic specific epistemic beliefs: Extending the theory of integrated domains in personal epistemology. *Learning and Instruction*, 56, 84–97.
- Molenaar, D. (2021). A flexible moderated factor analysis approach to test for measurement invariance across a continuous variable. *Psychological Methods*, 26(6), 660.
- Mueller, R. O., & Hancock, G. R. (2019). Structural equation modeling. In *The reviewer's guide to quantitative methods in the social sciences*, 2nd ed (pp. 445–456). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9781315755649-33>
- Nissani, M. (1996). The greenhouse effect: An interdisciplinary perspective. *Population and Environment*, 17(6), 459–489. <https://doi.org/10.1007/BF02208336>
- Oreskes, N. (2019). *Why Trust Science?* Princeton University Press.
- Pearl, J. (2023). The causal foundations of structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 49–75. Guilford.
- Plucker, J. A., & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist*, 56(2), 90–100. <https://doi.org/10.1080/00461520.2021.1895796>
- Preacher, K. J., & Yaremych, H. E. (2023). Model selection in structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 206–222. Guilford.
- Reeve, J., & Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist*, 56(1), 54–77. <https://doi.org/10.1080/00461520.2020.1862657>
- Reiss, J. (2015). A pragmatist theory of evidence. *Philosophy of Science*, 82(3), 341–362. <https://doi.org/10.1086/681643>
- Renkl, A. (2013). Why practice recommendations are important in use-inspired basic research and why too much caution is dysfunctional. *Educational Psychology Review*, 25(3), 317–324. <https://doi.org/10.1007/s10648-013-9236-0>
- Robinaugh, D. J., Haslbeck, J. M. B., Ryan, O., Fried, E. I., & Waldorp, L. J. (2021). Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction. *Perspectives on Psychological Science*, 16(4), 725–743. <https://doi.org/10.1177/1745691620974697>
- Robinson, D. H., Levin, J. R., Thomas, G. D., Pituch, K. A., & Vaughn, S. (2007). The incidence of “causal” statements in teaching-and-learning research journals. *American Educational Research Journal*, 44(2), 400–413. <https://doi.org/10.3102/0002831207302174>
- Robinson, D. H., Levin, J. R., Schraw, G., Patall, E. A., & Hunt, E. B. (2013). On going (way) beyond one's data: A proposal to restrict recommendations for practice in primary educational research journals. *Educational Psychology Review*, 25(2), 291–302. <https://doi.org/10.1007/s10648-013-9223-5>
- Sasco, A. J., Secretan, M. B., & Straif, K. (2004). Tobacco smoking and cancer: A brief review of recent epidemiological evidence. *Lung Cancer*, 45, S3–S9. <https://doi.org/10.1016/j.lungcan.2004.07.998>
- Schmidt, K., Merk, S., Rosman, T., Edelsbrunner, P. A., & Cramer, C. (in press). When perceived informativity is not enough: How teachers perceive and interpret statistical results of educational research. *Teaching and Teacher Education*.
- Schuberth, F. (2021). The Henseler-Ogasawara specification of composites in structural equation modeling: A tutorial. *Psychological Methods*. <https://doi.org/10.1037/met0000432>

- Seaton, M., Marsh, H. W., & Craven, R. G. (2009). Earning its place as a pan-human theory: Universality of the big-fish-little-pond effect across 41 culturally and economically diverse countries. *Journal of Educational Psychology, 101*(2), 403–419. <https://doi.org/10.1037/a0013838>
- Spearman, C. (1907). Demonstration of formulæ for true measurement of correlation. *The American Journal of Psychology, 18*(2), 161–169. <https://doi.org/10.2307/1412408>
- Spearman, C. (1927). *The abilities of man; their nature and measurement*. Macmillan Co.
- Thurstone, L. L. (1940). Current issues in factor analysis. *Psychological Bulletin, 37*, 189–236. <https://doi.org/10.1037/h0059402>
- Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*, 842–861. <https://doi.org/10.1037/0033-295X.113.4.842>
- Warne, R. T., & Burningham, C. (2019). Spearman's g found in 31 non-Western nations: Strong evidence that g is a universal phenomenon. *Psychological Bulletin, 145*(3), 237. <https://doi.org/10.1037/bul0000184>
- Werts, C. E., & Linn, R. L. (1970). Path analysis: Psychological examples. *Psychological Bulletin, 74*, 193–212. <https://doi.org/10.1037/h0029778>
- Werts, C. E., & Watley, D. J. (1969). A student's dilemma: Big fish-little pond or little fish-big pond. *Journal of Counseling Psychology, 16*(1), 14–19. <https://doi.org/10.1037/h0026689>
- West, S. G., Wu, W., McNeish, D., & Savord, A. (2023). Model fit in structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 184–205. Guilford.
- What Works Clearinghouse (2022). *Procedures and Standards Handbook, Version 5.0*. Institute of Education Sciences.
- Widaman, K. F., & Olivera-Aguilar, M. (2023). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (2nd Ed), pp. 367–384. Guilford.
- Wolff, F., Helm, F., & Möller, J. (2019). Integrating the 2I/E model into dimensional comparison theory: Towards a comprehensive comparison theory of academic self-concept formation. *Learning and Instruction, 62*, 64–75. <https://doi.org/10.1016/j.learninstruc.2019.05.007>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.