**EDITORIAL: The contribution of Multilevel Structural Equation Modelling to contemporary trends in educational research**

James Hall

University of Southampton, UK\*

Lars-Erik Malmberg

University of Oxford, UK

\* Corresponding Author. Southampton Education School, University of Southampton, UK, SO17 1BJ

Tel.: +44 (0)2380 593472 Email: J.E.Hall@Soton.ac.uk

**Keywords:** Multilevel Structural Equation Modelling; Structural Equation Modelling; Multilevel Modelling; Hierarchical Linear Modelling

**EDITORIAL: The contribution of Multilevel Structural Equation Modelling to contemporary trends in educational research**

**1. A brief history of Multilevel Structural Equation Modelling in educational research and a snapshot of today**

The potential contribution of Multilevel Structural Equation Modelling (MSEM[[1]](#footnote-1))to educational research was recognised early in the development of MSEM -- at least within Anglophone peer-reviewed journal papers[[2]](#footnote-2). Just four years after the first journal article to mention, ‘Multilevel Structural Equation Model’ (Lee, 1990), the first application of MSEM to educational research is encountered via, *‘…a multilevel path analysis on educational data in the Netherland’*  (Hox, Dronkers, and Schijf, 1994; p.45). Figure 1 shows the volume of publications indexed by Google Scholar per year to mention the term “Multilevel Structural Equation Model” since 1990. A sharp rise in publications mentioning the term can be seen from 2019 – the same year that the we proposed this Special Issue to the International Journal of Research and Method in Education and nearly thirty years after MSEM can be found first mentioned in a peer-reviewed English language academic journal publication.

**Figure 1.** The number of publications indexed by Google Scholar (as of 01/07/2020) that contain the term, “Multilevel Structural Equation Model” since 1990

Twenty six years on from Hox, Dronkers and Schijf (1994) and a reasonable question to ask might be, ‘To what extent has the potential contribution of MSEM to educational research been realised?’ While this question is open to debate, a snapshot of the presence of MSEM within the education journals associated with various professional research bodies concerned with educational research is indicative of the uptake of MSEM in this field. Table 1 provides this snapshot. It also includes three journals that are not affiliated with professional bodies (for comparison), and descriptively compares the presence of the term, ‘Multilevel Structural Equation Model’ over time against terms that speak to the origins of MSEM: ‘Structural Equation Model’ (SEM; excluding MSEM) and, "Multilevel Model" (or ‘Hierarchical Linear Model’; MLM/HLM).

As might be expected, Table 1 suggests a lag-effect between the development of MSEM (and its potential use in educational research) in the early 1990s and its emergence in peer-reviewed education journal articles (arguably still emerging after a slow start in the early 2000s). However by comparing the trends shown in Table 1 with that shown in Figure 1, we can observe that the slow uptake of MSEM research is not unique to education. It is only as of 2019 (twenty nine years after Lee, 1990) that we observe more than 200 publications a year featuring this term. In historical context, at the time of the MSEM papers of Lee (1990) and Hox, Dronkers, and Schijf (1994), a good number of the education journals shown in Table 1 were just starting to publish papers on either SEM or MLM/HLM[[3]](#footnote-3). Thus we observe a slow uptake of MSEM over the past thirty years – but an uptake that is now rapidly increasing.

This Special Issue of the International Journal of Research and Method in Education is (in part) a response to this increasing presence of MSEM in educational research (Table 1) and to MSEM in research more generally (Figure 1). It collects seven papers that use MSEM in different ways - to analyse data from empirical investigations and in the development of new methodological and analytical approaches. Before considering how these papers illustrate MSEM’s contribution to contemporary trends in educational research, and given the paucity of MSEM papers until recently, this editorial first presents a simple description of MSEM and then considers types of educational research for which MSEM has application. Both sections serve as primers: one to the analytic technique itself and one to its potential contributions to educational research (contemporary trends or otherwise). The intent is to equip readers with a foundational understanding of the technique and its application in order to allow them to more fully engage with MSEM - even if only conceptually (rather than statistically). This because readers of educational researchers are (demonstrably) at increasing likelihood of encountering MSEM (cf. Figure 1 and Table 1).

**Table 1.** The presence of “multilevel structural equation modelling” over time in education journals, psychology of education journals, and sociology of education journals

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Professional body** | **Journal title (education focussed)** | **Articles featuring "multilevel structural equation model"** | | | **Articles featuring "structural equation model" (excluding "multilevel")** | | | **Articles featuring "multilevel model" OR "hierarchical linear model"** | | |
| **n** | **earliest** | **latest (to 01/07/2020)** | **n** | **earliest** | **latest (to 01/07/2020)** | **n** | **earliest** | **latest (to 01/07/2020)** |
| American Educational Research Association | American Educational Research Journal (1964-) | 2 | 2017 | 2018 | 30 | 1978 | 2019 | 96 | 1989 | 2020 |
| Educational Evaluation and Policy Analysis (1979-) | 0 | - | - | 8 | 1983 | 2019 | 88 | 1990 | 2019 |
| Educational Researcher (1972-) | 0 | - | - | 19 | 1986 | 2020 | 39 | 1974 | 2020 |
| Journal of Educational and Behavioral Statistics (1976-) | 3 | 1997 | 2003 | 34 | 1995 | 2020 | 107 | 1994 | 2019 |
| Review of Educational Research (1931-) | 0 | - | - | 12 | 1980 | 2018 | 24 | 1994 | 2020 |
| Review of Research in Education (1973-) | 0 | - | - | 0 | - | - | 14 | 1988 | 2020 |
| American Psychological Association | Journal of Diversity in Higher Education (2008-) | 0 | - | - | 5 | 2016 | 2018 | 5 | 2008 | 2019 |
| Journal of Educational Psychology (1910-) | 18 | 2003 | 2019 | 194 | 1984 | 2020 | 162 | 1978 | 2020 |
| School Psychology (1986-) | 6 | 2013 | 2019 | 106 | 1992 | 2020 | 176 | 1998 | 2020 |
| Training and Education In Professional Psychology (2006-) | 0 | - | - | 8 | 2012 | 2020 | 3 | 2009 | 2017 |
| American Sociology Association | Sociology of Education (1927\*-) | 0 | - | - | 26 | 1977 | 2020 | 85 | 1985 | 2020 |
| British Educational Research Association | British Educational Research Journal (1975\*\*-) | 0 | - | - | 8 | 2007 | 2020 | 46 | 1991 | 2019 |
| British Psychology Society | British Journal of Educational Psychology (1931-) | 7 | 2015 | 2020 | 59 | 1992 | 2020 | 25 | 1999 | 2020 |
| - | Teaching and Teacher Education (1985-) | 3 | 2016 | 2018 | 53 | 1986 | 2019 | 47 | 1996 | 2020 |
| - | Learning and Instruction (1991-) | 5 | 2004 | 2018 | 51 | 1998 | 2020 | 52 | 1995 | 2020 |
| - | British Journal of Sociology of Education (1980-) | 0 | - | - | 4 | 2005 | 2020 | 7 | 1985 | 2019 |

*Notes: Publications as indexed by Google Scholar (as of 01/07/2020); The British Sociological Association publishes no journal specific to sociology of education although their website includes links to the British Journal of Sociology of Education; \*From 1927-1963 as, “The Journal of Educational Sociology”; \*\*From 1975-1977 as, “Research Intelligence”*

**2. Multilevel Structural Equation Modelling (MSEM) in a nutshell**

MSEM unites two statistical techniques that were independently developed in the early- to mid-20th century and which are today common within English language academic journals publishing educational research (see Table 1): Structural Equation Modelling (SEM)[[4]](#footnote-4) and Multilevel (or Hierarchal Linear) Modelling (MLM/HLM)[[5]](#footnote-5). MSEM therefore has historic roots in techniques including: biometric models (Wright, 1921), sociological models (Robinson, 1950; also referred to as ‘causal models’ e.g., Blalock, 1970), factor analysis (see Cudeck & MacCallum, 2007), and econometric models (see Matsueda, 2012).

For readers more familiar with SEM, MSEM extends these techniques via novel application to ‘nested data’ – data where cases are not independent from one another. In educational research this is typically when students are sampled from the same class/teacher/school (‘nested cases’) and/or when repeated measures (e.g. test scores) are gathered overtime from the same individuals (‘nested time points’). For readers more familiar with MLM/HLM (designed to analyse ‘nested data’), MSEM extends these techniques by integrating them with a range of existing techniques (from SEM) that relate to: measurement, classification, and models in which a concept can both be dependent and independent on others (e.g. when investigating mechanisms of effect via intermediate variables).

For readers unfamiliar with either SEM or MLM/HLM, both techniques are increasingly featuring in textbooks aimed at postgraduate (and sometimes undergraduate) instruction. For example, MLM/HLM was introduced in the 2009 (third) edition of, “Discovering Statistics using SPSS” (Field, 2009) – a textbook common on both undergraduate and postgraduate social science degrees. Meanwhile, MLM/HLM was also brought in to the 2007 (fifth) edition of, “Using Multivariate Statistics” (Tabachnick and Fidell, 2007) – a textbook common on postgraduate social science degrees and one that has featured a chapter on SEM by Ullman since the book’s 1996 (third) edition (Ullman, 1996). Perhaps tellingly, neither book yet includes a chapter on MSEM - although for interested readers, good primers on MSEM include the book by Heck and Thomas (2020) and the book chapters by Kaplan (2009) and Hox (2002).

One way to conceive of MSEM is to distinguish between two approaches to handling nested data within an SEM approach to statistical analysis: 1. An SEM where the standard errors of statistical estimates are adjusted to take into the non-independence of sampling into consideration, and 2. An MSEM characterised by one model per ‘level’ of analysis (e.g. one model for pupils and one for classrooms, or one model for occasions of measurement and one for individuals). The former can be regarded as a simple modification of SEM to make estimates of statistical probability robust to non-independence. Though fast to implement, this ‘flat’ SEM approach lacks the ability to answer research questions regarding measurement and relationships that may vary across levels (e.g. individual students versus whole classes) with the same level of detail as the MSEM alternative. For example, in answering research questions regarding whether patterns of change in student progress are consistent across classrooms, or the extent to which the mechanism of an educational intervention holds true across different schools. Of course, sometimes the simpler approach is more appropriate (and equally sometimes it is not). For readers familiar with MLM/HLM, this is reminiscent of the choice as to whether to statistically model not only random intercepts, but also random slopes, and sometimes also the correlation/covariance between random intercepts and slopes. This is a particularly pertinent comparison because examination of random slope and intercept effects (as conceived in MLM/HLM) requires the MSEM rather than the ‘flat’ SEM approach. In deciding between which of the approaches to use however, a balance must be struck between the parsimony and the appropriateness of two approaches as alternative options towards obtaining suitable and meaningful answers to given research questions. For a worked example of the two approaches used to answer the same research question see the paper by Bardach, Lüftenegger, and Yanagida (2020) in this special issue.

At the time of writing, MSEM innovation continues beyond just application to substantive fields (including education) but also within MSEM’s foundational mathematics. This is because disaggregated MSEM involves taking a suite of (SEM) statistical techniques developed for a single level (some with 100+ years of history such as factor analysis; Cudeck & MacCallum, 2007) and extending them to consider issues of consistency across time and/or across groups. Further, this requirement for continuing foundational development in mathematics comes not just because of the needs of MSEM, but because of continuing developments within SEM and MLM/HLM themselves. For example, in 2014 mathematicians had yet to develop a way to meaningfully incorporate samples weights within MLM/HLM (see Sammons et al., 2015). Thus, possible contributions of MSEM to educational research, first acknowledged in a journal article twenty six years ago and now increasingly recognised, are not bounded by what is technically possible today, and are not bounded by the examples contained in this Special Issue. Rather, today’s educational researcher has a suite of MSEM tools at their disposal that others are continuing to improve and develop. Mapping the uses of these tools to different areas of education fosters their use, fosters understanding of their use, and therefore fosters the contribution of MSEM to trends both present and future.

**3. Educational research designs and questions for which MSEM has application**

Table 2 shows that MSEM has application across both experimental and non-experimental research designs plus application to studies where there are comparisons of multiple groups (creating nested cases), comparisons of multiple time points, or both. This broad applicability of MSEM means that it is not limited in application to any type of research design (e.g. experiments, surveys, case studies) or to any type of research method/instrument (e.g. questionnaires, interviews, observations, tests [educational or psychological]). Instead, MSEM can be carried out on any data for which MLM/MLM is suitable. Why then would one use MSEM instead of MLM/HLM? Following on from the previous section, this distinction relates to the nature of the research question being investigated.

First, MSEM has unique application to research questions relating to variation in measurement and classification across groups and/or time. Where there are unobserved constructs to be measured (e.g. via one of the many types of factor analyses, IRT, Rasch modelling) or latent groups to be identified (e.g. via mixture modelling; see Vermunt, 2008) then MSEM provides a statistical framework to explore the extent to which that ‘latent’ measurement and/or ‘latent’ group identification is consistent over groups and/or time (e.g. Asparouhov and Muthén, 2007). For example, does the ‘latent’ psychological construct of student self-efficacy remain conceptually equivalent over time and/or across students in different countries? Alternatively, to what extent is the composition of a group of students identified as ‘disadvantaged’ consistent across different schools?

**Table 2.** Educational research designs for which MSEM has application – *note the 100% overlap with MLM/HLM in terms of application to research designs*

|  |  |  |
| --- | --- | --- |
| **Research Design Type** | **Nested Data Type** | **Example Research Designs** |
| Experimental (primary or secondary) | Nested Cases | True- or Quasi-Experimental study of an student intervention rolled out in multiple classrooms or schools |
| Nested Time points | Experiment that considers only change over time in a group after a change introduced by the researchers |
| Both | True- or Quasi-Experimental study of an intervention rolled out in multiple classrooms or schools - with repeated measurement of the dependent variable(s) |
| Non-Experimental (primary or secondary) | Nested Cases | Survey or Case Study that compares naturally occurring groups (including groups that are compared in a natural experiment) |
| Nested Time points | Longitudinal Cohort Study |
| Both | Longitudinal Cohort Study that achieved nested cases via cluster sampling (e.g. children in preschools) |

Second, MSEM has unique application to a number of research questions relating to models in which a concept can both be dependent and independent on other (see Christ et al., 2017) – for example to research questions postulating mediation effects and/or indirect effects (e.g. Preacher, Zyphur, and Zhang, 2010). Again, when such questions consider variation across groups and/or time then MSEM can have unique applications that are unavailable to other statistical approaches. Moreover, such unique applications of MSEM can be combined with those mentioned in the paragraph above. This is illustrated by extending the example research questions just considered. For example, does student self-efficacy mediate the relationship between teacher effectiveness and student attainment and if so, how does this differ over time and/or across countries?

MSEM provides a single analytic framework for answering all of the above example questions. This common framework has a number of practical benefits for research and for researchers (education-focussed or otherwise) including: 1. Statistically-compatible analyses. The results of one study can be extended in another at the statistical rather than the conceptual level (thus increasing validity); 2. Reduced requirement to learn and maintain working-knowledge of multiple statistical analyses and of multiple statistical software programs; 3. Adaptability to future trends. As a broad suite of techniques, SEM is being continually developed and expanded by academics around the world. Thus, as new statistical techniques are developed they can be rapidly integrated both within the statistical framework of SEM and its computer software (as is currently being experienced in the case of Bayesian SEM; see Section 5).

**4. Seven illustrations showing how MSEM can contribute to contemporary trends in educational research**

Having described the history and prevalence of MSEM, outlined what it is, and explained types of research questions for which MSEM has application beyond MLM/HLM, this section presents a snapshot of the different ways in which this analytic approach can contribute to contemporary trends in educational research. This is achieved by considering each of the seven papers featured in this special issue and for each drawing out the area of educational research that the paper is concerned with and highlighting the contribution made by MSEM. Both nesting of students in classes/schools and nesting of time points in students can be found in these papers - as can both the types of research question considered in Section 3.

Beginning with Bardach, Lüftenegger, and Yanagida (2020), this paper focusses on how the group psychosocial processes and goal structures in a classroom (the ‘classroom climate’) can influence the motivation of individual students. A comparison of the cluster-robust standard error approach (see Section 2) and a ‘doubly latent’ MSEM approach for statistically modelling the nesting of students within classrooms is undertaken with data from 1645 Austrian secondary school students in 83 classrooms. MSEM is used to explore differences in the measurement and relationships between three psychological constructs (mastery goal structure, social goal structure, personal mastery goals) as they exist both for individual students and for whole classrooms. Empirical results favour use of the more complex approach in studies investigating the effects of classroom climate.

Second, Burns, Martin, and Collie (2020) consider the school-level effects of students’ academic achievement and compare the effects of perceived ‘school belonging’ and perceived teacher unfairness – as experienced both by individual students and by schools as a whole. Data to compare the effects of these perceptions on student achievement comes from the 2015 PISA dataset (210,381 students in 7964 schools from 32 OECD countries). Again, MSEM is used investigate differences in the measurement and relationships between constructs as they exist both for individual students and as composite scores for whole schools. The effects of perceived ‘school’ belonging’ and teacher unfairness on students’ achievement are different depending on whether these perceptions are considered across a whole school on for individual students. In turn, these suggest that there are distinct psychosocial processes at play at each level.

Third, Hall et al. (2020). This paper is directly concerned with the trend for increasing MSEM in educational research that has been earlier outlined in this editorial. MSEM is used to reconsider the definition and implementation of a fundamental concept in contemporary multivariate statistical analysis in the social sciences (Moderation) and by doing so identifies an under-explored extension that the authors term “Airbag Moderation”. The role of MSEM is critical in this paper. It is through MSEM that the core assumption adhered to in contemporary statistical analysis is challenged: That Moderation is conceptually equivalent to Interaction due to both being able to be statistically implemented via statistical interaction terms. As MSEM has developed its focus on random slope effects (see Section 2) this has led to the development of latent variable approaches to represent these effects – an approach that can be implemented in any data set – not just those with nested cases or nested time points. Because random slope effects allow for the testing of Moderation without use of statistical interaction terms, so this new development permits the assumption to be challenged that Moderation is the same as Interaction (in part due to the fact that both can be statistically tested with statistical interaction terms). That then opens up the opportunity for new methodological and statistical concepts to be developed, of which Airbag Moderation is one. Data from 2568 English children aged 1-3 years who used one of 117 Sure Start Children’s Centres is then analysed to demonstrate that Children’s Centres function in this manner in the context of poverty and internalising behaviours shown by 3 year olds.

Fouth, Kyriakides and Charalambous (2020) provide an explicit illustration of the importance of MSEM to theory testing within a hypothetico-deductive framework of knowledge acquisition within educational research. This is achieved by considering both direct and indirect effects on student learning outcomes that can result from whole school interventions (here the Dynamic Approach to School Improvement, DASI; see Creemers and Kyriakides, 2010). Data come from 5560 students (Grades 4-6) in 72 primary schools from four European countries (Cyprus, England, Greece, Ireland). MSEM was used to estimate both direct and indirect effects (see Section 3) of the DASI intervention on student achievement in mathematics.

Fifth, the paper by Malmberg (2020) responds to a growing trend for educational research that is focussed on learning processes with a didactic paper that explains and demonstrates the utility of an MSEM that is suitable for use with data where there are multiple time points nested within cases (Intraindividual Structural Equation Modelling, ISEM). The paper provides step-by-step instructions on how to use ISEM to analyse process data and this is followed by an empirical demonstration. This demonstration features real-time educational data from a longitudinal study of 202 primary school students (Grades 5 and 6) in England plus measures of student’s self-reported motivation to learn (reported 10-24 times by each child) and teacher expectations. The ISEM involves using ‘individuals as their own controls’ and takes account of the chronological order of the obtained data.

Sixth, the paper by Marcoulides and Yuan (2020) focusses on a universal aspect of every MSEM: How can researchers demonstrate that their theory-driven sets of relationships (their statistical models) are appropriate for the data to which they are applied (AKA: model fit)? Marcoulides presents a new approach for testing model fit that is based on extending an approach currently used in single-level (‘flat’) SEM (equivalence testing). The benefits of using this approach in MSEM are illustrated via estimation of a multilevel Confirmatory Factor Analysis (CFA) of synthetic data representing responses by 421 teachers from 31 primary schools to six items measuring ‘socio-curricular structure’ (see Marcoulides, Heck, and Papanastasiou, 2005). Comparing the equivalence testing approach to alternatives (currently used in MSEM), results favour the new approach.

Seventh, Pakarinen, Lerkkanen, and Suchodoletz (2020) look at the emotional support that teachers provide to preschoolers and the association that this has with their developing social competence. Data came from 47 qualified teachers in Finland and their classrooms within which were 447 children aged rising 6 years from autumn to spring. MSEM is used to investigate reciprocal effects (auto-regressive and cross-lagged associations) between teachers’ emotional support and children’s pro-social and anti-social behaviour at both the individual and classroom level. Greater emotional support is associated with improvements in social competence over the academic year.

**5. Looking ahead: Realising the potential of MSEM for educational research and signposting challenges**

Despite the utility of MSEM for educational research being known for twenty six years, the presence of MSEM within educational journals has been slow to take off but is now accelerating (just as MSEM is across all fields; see Figures 1 and 2) and there are practical reasons for this and practical ways in which this pace may be increased in order for the potential utility of MSEM to educational research to be achieved. Considering first reasons for the surge of MSEM in journal articles within recent years, a number of recent developments are likely to have helped and these include the MSEM packages now available within the freeware statistical software “R”, and the emergence of text primers on MSEM (see Section 2). There is also a demonstrable culture change happening as regards the educational research that is engaged in by journal authors, editors, and reviewers (given the publication of education journal articles featuring MSEM): All three are increasingly willing to engage in the writing, reading, and consideration of papers that use this technique. The increasing presence of MSEM in educational journals would not be possible unless this has taken place. That said, if MLM/HLM and SEM can be regarded as ‘mainstream’ statistical techniques in contemporary educational research (Table 1), then perhaps only when MSEM reaches equivalent presence across education journals may it be reasonably regarded in the same terms.

Practical ways in which the pace of take-up of MSEM might be increased (again, to help it fulfil its potential utility to educational research) include: more easily accessed statistical software, more easily accessed training (postgraduate and continuous professional development), and increased cultural support for the use of this technique across the varied groups of researchers that generate educational research. Prior to the availability of the R freeware package the costs of commercial SEM software (particularly in the developing world) could only function as an impediment to the use of MSEM where it was appropriate. However, alongside the availability of MSEM packages in R, MSEM training is becoming increasingly available and is facilitated by online communities, courses, and texts (e.g. Heck and Thomas, 2020). However, the extent to which this training is taken up will be constrained or enabled by the professional cultures within which this training is undertaken (as with the use of MSEM in educational research). At the time of writing and from Figure 1, MSEM seems to have gained particular traction within research concerning the psychology of education, but whether (and when) that will be replicated in other areas of educational research remains to be seen.

In terms of technical challenges impeding the take-up of MSEM in research (education-focused or otherwise), current issues likely to continue include difficulties handling cross-classified data structures (or data structures where participants switch between groups over time; e.g. changing schools in a longitudinal study), obtaining meaningful estimates of model fit in MSEM, and the creation of flexible tools for accurate estimations of statistical power in circumstances where there are unequal cases per group or time points per case. While Bayesian approaches to MSEM can help with a number of these[[6]](#footnote-6), switching to this approach poses challenges both for the underlying SEM techniques (e.g. for assessing model fit; Garnier-Villarreal and Jorgensen, 2020) and risks slowing the pace at which MSEM is taken up (when required; see Section 3) by educational research due to the extra learning that is required to access not only MSEM but also to the Bayesian approach to statistics[[7]](#footnote-7).

**References**

Asparouhov, T., and Muthén, B.O. (2007). Multilevel Mixture Models. In G.R. Hancock and K.M. Samuelsen (eds.)*, Advances in Latent Variable Mixture Models.* Charlotte, NC: Information Age Publishing

Asparouhov, T., and Muthén, B.O. (2012). General Random Effect Latent Variable Modeling: Random Subjects, Items, Contexts, and Parameters. *Annual Meeting of the National Council on Measurement in Education*, Vancouver, British Columbia

Bardach, L. , Lüftenegger, M. and Yanagida, T. 2020. Studying Classroom Climate Effects in the Context of Multi-level Structural Equation Modelling: An Application-Focused Theoretical Discussion and Empirical Demonstration. *International Journal of Research & Methods in Education, 43*(4). https://doi.org/10.1080/1743727X.2020.1791071

Blalock, H.M. (1970). Estimating Measurement Error Using Multiple Indicators and Several Points in Time. *American Sociological Review, 35*(1), 101-111. Retrieved from: http://www.jstor.com/stable/2093857

Burns, E.C., Martin, A.J., and Collie, R.J. (2020). Supporting and Thwarting Interpersonal Dynamics and Student Achievement: A Multi-Level Examination of PISA 2015, *International Journal of Research & Method in Education*, *43* (4). https://doi.org/10.1080/1743727X.2020.1757639

Christ, O., Hewstone, M., Schmid, K., Green, E.G.T., Sarrasin, O., Gollwitzer, M., and Wagner, U. (2017). Advanced Multilevel Modeling for a Science of Groups: A Short Primer on Multilevel Structural Equation Modeling. *Group Dynamics: Theory, Research, and Practice, 21*(3), 121–134. https://doi.org/10.1037/gdn0000065

Coffman, W.E., and Lee, L.M.P. (1974). Cross-National Assessment of Educational Achievement: A Review. *Educational Researcher, 3*(6), 13-16. https://doi.org/10.2307/1175234

Creemers, B.P., & Kyriakides, L. (2010). Using the Dynamic Model to Develop an Evidence-Based and Theory-Driven Approach to School Improvement. *Irish Educational Studies, 29*(1), 5-23. https://doi.org/10.1080/03323310903522669

Cudeck, R., and MacCallum, R.C. (2007). *Factor Analysis at 100: Historical Developments and Future Directions.* Mahwah, NJ: Laurence Erlbaum

Field, A. (2009). *Discovering Statistics using SPSS (and Sex and Drugs and Rock'n'Roll)* (3rd ed.). Londnon, UK: Sage

Garnier-Villarreal, M., and Jorgensen, T.D. (2020). Adapting Fit Indices for Bayesian Structural Equation Modeling: Comparison to Maximum Likelihood. *Psychological Methods, 25*(1), 46–70. https://doi.org/10.1037/met0000224

Goldstein, H. (2011). *Multilevel Statistical Models* (4th ed.). Chichester, UK: John Wiley & Sons.

Hall, J., Malmberg, L-E., Lindorff, A., Baumann, N., and Sammons, P. (2020). Airbag Moderation: The Definition and Implementation of a New Methodological Model.  *International Journal of Research & Method in Education*, .*43*(4). <https://doi.org/10.1080/1743727X.2020.1735334>

Heck, R.H., and Thomas, S.L. (2020). *An Introduction to Multilevel Modeling Techniques: MLM and SEM Approaches using Mplus* (4th ed.). New York, NY: Routledge

Hox, J.J., Dronkers, J., and Schijf, H. (1994).An Application of a Multilevel Structural Equation Model. In F. Geyer, K.M. van Meter, and R. Pawson. *Research Committee on Logic and Methodology (RC33): Research Committee 33 of the International Sociological Association: RC33 Newsletter Spring 1994, 42*. International Sociological Association. Retrieved from: <https://search.proquest.com/docview/61727695?accountid=13963>

Hox, J.J. (2002). *Multilevel Analysis. Techniques and Applications.* Mahwah, NJ: Lawrence Erlbaum

Kaplan, D. (2009). Multilevel Structural Equation Modeling. In D.Kaplan. (ed.), *Structural Equation Modeling: Foundations and Extensions* (2nd ed.,), 133–154. Los Angeles, CA: Sage

Kyriakides, L., and Charalambous, E. (2020)*.* Testing the Theoretical Framework of Whole School Interventions Aiming to Promote Student Learning Outcomes: The Contribution of Multilevel Structural Equation Modelling. *International Journal of Research & Method in Education, 43*(4). https://doi.org/10.1080/1743727X.2020.1795114

Lee, S.Y. (1990). Multilevel Analysis of Structural Equation Models. *Biometrika*, *77*(4), 763-772.<https://doi.org/10.1093/biomet/77.4.763>

Lee, S.Y. (2007). *Structural Equation Modeling: A Bayesian Approach.* New York. NY: John Wiley & Sons

Malmberg, L-E. (2020). Intraindividual Structural Equation Models for Learning Experiences. *International Journal of Research & Methods in Education*, *43*(4). <https://doi.org/10.1080/1743727X.2020.1793939>

Marcoulides, G.A., Heck, R.H., and Papanastasiou, C. (2005). Student Perceptions of School Culture and Achievement: testing the Invariance of a Model. *International Journal of Educational Management, 19*, 140-152. https://doi.org/10.1108/09513540510582435

Marcoulides, K.M., and Yuan, K-H.  (2020).  Using Equivalence Testing to Evaluate Goodness of Fit in Multilevel Structural Equation Models. *International Journal of Research & Method in Education, 43*(4). https://doi.org/10.1080/1743727X.2020.1795113

Matsueda, R.L. (2012). Key Advances in the History of Structural Equation Modeling. In R.H. Hoyle (ed.), *Handbook of Structural Equation Modeling,* 17–42. New York, NY: Guildford Press

Pakarinen, E., Lerkkanen, M.-K. and Suchodoletz, A. (2020). Teacher Emotional Support in Relation to Social Competence in Preschool Classrooms. *International Journal of Research & Method in Education 43*(4). https://doi.org/10.1080/1743727X.2020.1791815

Preacher, K.J., Zyphur, M.J., and Zhang, Z. (2010). A General Multilevel SEM Framework for Assessing Multilevel Mediation. *Psychological Methods, 15*(3), 209. https://doi.org/10.1037/a0020141

Robinson, W.S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review, 15* (3), 351–357. https://doi.org/10.1093/ije/dyn357

Sammons, P., Hall, J., Smees, R., Goff, J., Sylva, K., Smith, T., Evangelou, M., Eisenstadt, N., and Smith, G. (2015)*.* *Evaluation of Children’s Centres in England (ECCE). Strand 4: The Impact of Children’s Centres: Studying the Effects of Children’s Centres in Promoting better Outcomes for Young Children and their Families.* Research Report DFE-RR495. London, UK: DfE. Retrieved from: https://ore.exeter.ac.uk/repository/bitstream/handle/10871/23328/DFE-RR495%20Strand%204%20Research%20Report%20(Sammons%20et%20al.,%202015).pdf?sequence=1

Tabachnick, B.G., and Fidell, L.S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Pearson

Tarka, P. (2018). An Overview of Structural Equation Modeling: Its Beginnings, Historical Development, Usefulness and Controversies in the Social Sciences. *Quality & Quantity, 52*(1), 313-354. https://doi.org/10.1007/s11135-017-0469-8

Ullman, J.B. (1996). Structural Equation Modeling. In B.G. Tabachnick and L.S. Fidell (eds.), *Using Multivariate Statistics* (3rd ed.), 709-811. New York, NY: Harper Collins

Vermunt, J.K. (2008). Latent Class and Finite Mixture Models for Multilevel Data Sets. *Statistical Methods in Medical Research, 17*(1), 33-51. https://doi.org/10.1177/0962280207081238

Wright, S. (1921). Correlation and Causation. *Journal of Agricultural Research*, *20*, 557-585

1. Sometimes also found represented with the acronym ‘ML-SEM’. Note also that ‘MSEM’ is sometimes instead used for, “Multi-group Structural Equation Model”. [↑](#footnote-ref-1)
2. By searching for the term, “Multilevel Structural Equation Model” within the electronic databases: ERIC, Google Scholar, PsychINFO, Social Services Abstracts, Sociological Abstracts, and Web of Science. [↑](#footnote-ref-2)
3. While keeping in mind the existence of far earlier publications – the earliest (to the best of our knowledge) being for MLM/HLM via Coffman and Lee (1974). Although meaning of the phrase “hierarchical linear model” in this paper (p.15) is ambiguous. The authors may not be referring to MLM/HLM as commonly used today but instead to a form of ‘flat’ multiple regression analysis. [↑](#footnote-ref-3)
4. For a history of SEM plus its uses and controversies in the social sciences see Tarka (2018). [↑](#footnote-ref-4)
5. For a primer on MLM/HLM see Goldstein (2011). [↑](#footnote-ref-5)
6. For example, in cases where is low sample size at higher levels of analysis (e.g. many pupils but relatively few schools), reduction of inaccessible parameter estimates, and for cross-classified modelling (e.g. Asparouhov and Muthén, 2012). [↑](#footnote-ref-6)
7. For interested readers, an award-winning primer on Bayesian approaches to SEM is the textbook of Lee (2007). [↑](#footnote-ref-7)