Towards a Logical Framework for Latent Variable Modelling

Trisha Nowland, Alissa Beath and Simon Boag

May 31, 2018

Towards a Logical Framework for Latent Variable Modelling

Trisha Nowland
Macquarie University
North Ryde
Australia
trisha.nowland@mq.edu.au

Alissa Beath
Macquarie University
North Ryde
Australia
alissa.beath@mq.edu.au

Simon Boag
Macquarie University
North Ryde
Australia
simon.boag@mq.edu.au

Abstract

Modelling in psychometrics has become increasingly reliant on computer software; at the same time, many decisions that a researcher makes remain unrecorded and perhaps, unreconciled to anything more than the researcher's intuition or best guess. The aim of this paper is to set out a logic that accounts for and guides decision procedures, in psychometric research practices. Such a logic is informed by the integration of three systematic viewpoints: i) bounded rationality (Gigerenzer & Goldstein, 1996); ii) axiomatic measurement theory (Suppes, 2002, 2006, 2016); and iii) a constructive mathematics (Ferreiros, 2016). The integration of these three systems under an overall perspective that is characterised by inference from the best systematisation (Rescher, 2016) is reviewed, and compared to current research practices, with particular reference to the problems for psychometric sciences that are revealed in the Reproducibility Project (Open Science Collaboration, 2012) outcomes. Our conclusion for the overall system is that the constraints characterised by constructive mathematics offer unique tools to the researcher in accounting for their decision procedures, and a proposal for a software tool that handles the decision protocol is presented. In characterising the decision procedures, we also explore the way that rough set theory (Pawlak, 1982) is integrated into decision procedures to provide insight into database fields or variables that hold some import but may otherwise remain hidden in research outcome reporting.

Keywords Constraints, constructivism, systematicity

1 Introduction

Psychometrics, broadly defined, is the scientific practice of utilizing theories, models, and instruments with the aim of measuring psychological phenomena (McDonald, 1999). First described by Galton in 1879 as "[t]he art of imposing measurement and number upon operations of the mind" (Galton, 1879, p.149), an important and much contested aspect of psychometric practice through its history has been that it involves a quantitative aspect of data as relevant to psychological phenomena (c.f Michell, 1997, 2003; Luce, 1997; Borsboom & Mellenbergh, 2004; Borsboom, 2006). The contest comes about in part because psychological phenomena are difficult to measure directly – that is, there is little direct or tangible empirical evidence regarding the state of affairs that persists in minds (Borsboom, 2008; Sijtsma, 2006). Statistical techniques have featured since the dawn of psychometrics as tools that notionally support the provision of insights into the operations,

of minds. One technique features the latent variable model, which comes into being in its earliest form in the work of Charles Spearman (1904) on his factor theory. The paper "*General Intelligence – Objectively Determined and Measured*" (1904) brings together analyses of academic test results from several studies of groups of students of different ages. Spearman had observed the positive manifold, the idea that students with high scores in one school study area tended to also have high scores in other areas. The correlational analysis outcomes revealed hierarchical correlation patterns. Based on this, he claimed that all test scores could be broken down into two parts. One part was a latent variable, which accounted for the common factor in the correlations for student's test scores. The other part was a unique element or factor that was the part of the outcome specific to that particular type of test only (Spearman, 1904). In this theory, the latent variable model was born.

1.1 Aim

The latent variable model was subject to substantial critique quickly after it first emerged (Thomson, 1916; Thorndike, 1927; Thurstone, 1934). This contributed to the proliferation of alternate although similar versions of the same factor theory. Some examples of variations included increasing the number of latent variables (Thurstone, 1934), or the ways in which they were considered to be related (Thomson, 1916). The computer revolution of the 1950s brought with it an array of solutions for latent variable modelling which was otherwise prohibitively labour intensive for the purposes of psychometric analysis. It also brought with it more problems connected to the increasing numbers of statistical assumptions, some of which are quite abstract but necessary for the utilization of the statistical techniques (see Cliff, 1983; Bentler & Chou, 1987). The aim of this paper is to set out some of the logical concerns that remain for us today with respect to the use of the latent variable model in psychometric practice. We maintain focus on the problem of factor indeterminacy, exploring the degree to which a framework grounded in logic provides some structure with which problems can be clarified, and resolved. A secondary intent is supporting rigorous and sustainable research practices underscored by a framework founded in principles informed by logical consistency, noting that psychological research itself increasingly references qualitative inputs, even in the process of producing quantitative outputs.

2 Latent Variable Model

2.1 What it is

A generalised definition of a latent variable model is a statistical model that specifies relationships between two types of variables, manifest variables and latent variables (Bollen, 2002; Maraun & Gabriel, 2013). The latent variable model in its most basic form is made up of at least three elements: i) the latent variable, often described as an underlying or hidden common factor; ii) the manifest variables, which are the actual test score outcomes or other data amongst which commonality can be expected; and iii) the relationships between the latent variable and each of the manifest variables, represented by paths (Skrondal & Rabe-Hesketh, 2004). An example from Spearman (1904) includes the common factor of general intelligence underpinning student manifest test score outcomes on tests for mathematics, language, classics, and so on. Any single latent variable demands for its evidence correlation or covariances from two or more manifest variables. In its initial form as elaborated by Spearman (1904) there was no inclusion of an error term for the variables or the relation between them. Subsequent developments that have included error terms and different variable structures as developed over the next 100 years include item response theory (Guttman, 1950), exploratory factor analysis (Jöreskog, 1971), confirmatory factor analysis (Jöreskog, 1971), structural equation modelling (Jöreskog, 1973), and latent profile analysis (Bartholomew, 1987), among others. Applications of the latent variable model in research are diverse, extending to personality theory (e.g. NEO-PI-R, Costa & McCrae, 2008), intelligence testing (e.g. WISC-R: Kaufman, 1979), quality of life scales (Fayers & Hand, 1997) and meta-analyses of clinical trials (Eusebi, Reitsma & Vermunt, 2014).

2.2 Model assumptions and problems

The latent variable model makes one core assumption, that of *conditional independence* (Lord & Novick, 1968). The assumption of conditional independence gives us the relation between the manifest variables and the latent variable, and in bare structure states that the only way that any correlation between the manifest variable outcomes can occur is due to the existence of the latent variable, or common underlying factor (Maraun, 1996). This is a strong assumption and one that is by necessity involved with the use of the model – without conditional independence there is no latent variable, in the model. As mentioned above secondary assumptions were added later in statistical developments that accompanies analyses using latent variable modelling. Several conceptual concerns are considered in the literature, including as examples questions of causality (Cliff, 1983), or the relevance or otherwise of correlations and covariances calculated on between-subjects data to any single individual (Borsboom, 2005; Weinberger, 2015). Further problems that remain for applications of the latent variable model that are not dealt with in this paper include decisions regarding calculation techniques for the parameters of the latent variable model, measurement invariance (Meredith, 1993), unidimensionality (McDonald, 1999), model identification (Romeijn & Williamson, 2018), and the problem of equivalent models (Maccallum et al., 1993).

2.3 Focus on conditional independence

In recent psychometric literature the conditional independence assumption has been used to claim that psychometricians who use the latent variable model must be realists, as the conditional independence assumption stipulates that some latent variable must exist (Borsboom, Mellenbergh, & van Heerden, 2003; Borsboom, 2005). This claim for realism extends to a claim for ontological and causal import for the latent variable – that is, the latent variable *causes* test score outcomes. The claim for a causal ontology for latent variable modelling follows from an analysis of latent variable modelling practice. It takes this practice as an object of analysis, and asks what sort of methodological commitment must be necessary, to use the latent variable model in the way that psychology researchers use it. The scientific realism of Hacking (1983) lays claim to an idea that theoretical entities must exist, before we can begin analyzing them. If analysis is taking place, there must be real theoretical entities, that we are concerned about. This account of scientific realism is used in Borsboom (2005) to suggest that latent variables must exist, because researchers go about investigating them in their analyses. This pragmatic perspective however leaves us with no ontological resources such as resources that we could use to ask question about the existence of the latent variable, or the phenomenon that the latent variable represents. There are no criteria made available to facilitate critique of the practices, or the claims about the latent variable. A logical realist ontology, distinct from a scientific realist ontology, for example, may be used to help ask questions about the existence of the phenomena, prior to the existence or otherwise of the latent variable, in a process of critical inquiry (Petocz & Newbery, 2010; Michell, 2000). More will follow regarding the potential for a logical ontology and a logical, but pragmatic, realism, below.

2.4 Factor indeterminacy

One problem that follows from the assumption of conditional independence is the problem of factor indeterminacy (Maraun, 1996; Guttman, 1955). The problem of factor indeterminacy is a mathematical problem inextricably connected to the conditional independence assumption that founds the model (Maraun, 1996). It says that given the nature of the relationships used as evidence for the latent variable, and the conditional independence assumption that simply states that the latent variable must exist, we can never with full certainty connect the outcomes of our analysis to the psychological attribute or construct that the latent variable is given to stand for (Mulaik, 2010). This problem was pointed out to Charles Spearman relatively early in the development of his factor theory (Wilson, 1928), although evidence in the literature point to the likelihood that Spearman misunderstood the nature of the problem (see Spearman, 1929; Mulaik, 2010). One way to describe the problem is to say that while we may locate our latent variable as one existing in a domain of possible variables, there is certain no way to connect any outcomes of our analysis to this particular latent variable – logically it could be this or any other latent variable within the domain (Guttman, 1955; Mulaik, 2010). Some suggestion occurs in the literature to state that one way to close the gap is to clarify the nature of what is described as an 'infinite behavior domain" (McDonald, 2003; Guttman, 1955). Such a domain allows the psychometrician to derive say the particular set of questionnaire items, where the scores of these items constitutive for the manifest variables – for example, as a summed test score. An assumption about an infinite structure for the domain allows psychometricians to obtain some limiting properties for the latent variable model, and to produce some outcomes in latent variable analysis conditioned on the assumption of infinite structure for the domain (McDonald, 2003). Infinite structure remains as an assumption, however, and infinity enters into many of the statistical practices that make up the different techniques of latent variable modelling. Nowhere in latent variable modelling practices is a researcher required to account for such assumptions. This holds true for latent variable modelling, and also for psychometric techniques, generally. The core purpose of the conceptual

framework as described below is to invite researchers to explicitly account for their research decisions and assumptions. In this way, the logical structure of any researcher's conclusions and associated risk levels can be scrutinized by fellow and future practitioners.

3 Logical framework

The concerns for the latent variable model as discussed above point to the need for a framework which facilitates scrutiny of researcher assumptions and decisions. This becomes important when we realise that Spearman (1904) tried but was never able to offer a mathematical proof for the original factor theory (Cowles, 2001). Further, the maximum likelihood technique, typically used to estimate parameters in the latent variable model, remained without mathematical proof despite the best efforts of its original developer, Ronald Fisher (see Stigler, 2007). The latent variable model therefore cannot be described offering deductive closure as a model, given these proof gaps (see Suppes, 1999; Tarski, 1930). Because of the gaps, use of the model itself cannot be taken as conclusive evidence for the existence of any latent variable. We require evidence beyond the model analysis itself for the existence of the latent variable, or perhaps more correctly, the psychological phenomenon represented by the latent variable. A logical conceptual framework can facilitate discernment of logical relations between elements of the research project, such that the logical consistency of the project can be evaluated. Such a framework supports the clarification of constraints implied by each aspect or domain of the research project in conceptualization and planning stages. What results is an understanding of the broader context of the evidence base, and the logical conditions for the psychological phenomenon in question. Evidence is derived in a sequence of logical relations (Petocz & Newbery, 2010) that are relevant to the phenomena, providing information beyond just the model itself. As in the approaches to constructive mathematics most clearly articulated by Ferrerios (2016), and evident in the advocacy for axiomatic approaches in Maddy (2011) and Suppes (2002), constraints form guiding principles for reasoning about phenomena. These constraints allow us to understand something of the nature of risk associated with any claim that follows from the completion of research analysis, given what the constraints indicate regarding what can and cannot be said. A logical framework providing a systemic view of the research project has a number of roles, including: i) guiding inquiry within the same research project; ii) elaborating a set of conditions or criteria for the research project for the evaluation of research outcomes; iii) allowing for a logical matching or testing of relations between the different aspects of research projects. The domains for a research project may include researcher stance; prevailing ideology; relevant theory and models; useful methodology; variables and their structures; data structures, and, the original phenomena at the heart of the research interest. We will maintain focus in what follows on points ii) and iii) above, and begin by exploring the relevance of the concept of inference from best systematization, for psychometric practice.

3.1 Inference from the best systematization

Acknowledging the limits of the hypothetico-deductive and deductive-nomological methods of evaluating scientific knowledge with their positivist roots (Hempel, 1965; Salmon, 1990), inference to the best explanation has recently been proposed as more relevant to present day practices that include the use of psychometric

techniques (Haig, 2009). Inference to the best explanation (IBE: Harman, 1979) is proposed as stepping away from logical deduction to look to evaluation against explanatory criteria for any proposed theory within psychology, aiming to maximize explanatory power with a view to claims of scientific truth (Haig, 2009). An initial question arises with respect to the explanatory aspect of IBE. Use of the latent variable model may be involved in any of the activities of description, explanation, or prediction that together make up scientific activity and which differ from each other in nature (see Boag, 2011). Rescher (2016) notes several gaps when we rely in IBE, an approach that implies we have been able to evaluate upfront what the best explanation actually is. There are no resources by which we can assess the tightness of reasoning, the soundness of outcomes in light of previous scientific knowledge, the generality of applicability of the outcomes across real-world situations, and so on. Inference from the best systematization (IBS: Rescher, 2016) invites the use of logical criteria against which we can evaluate the systematicity of a research project, in its earliest phase. A systematic approach invites scrutiny of logical links, appreciation of theoretical and methodological constraints, and articulation of possible gaps in our research program, before we begin any work. In this way IBS demands something like the processes which have now become commonplace in psychological research, where many of the psychometric assumptions and commitments are lodged in research pre-registration databases (see Open Science Collaboration, 2015). Such databases aimed originally at the question of evaluating reproducibility of research outcomes in the psychological sciences; a by-product has been more transparency regarding psychometric assumptions and practices. However, these pre-registration databases do not presently invite logical connection between aspects of research projects and evaluation of overall project risk. The logical framework proposed here for psychometric practices does exactly these things.

3.2 Bounded rationality and logic

IBS requires attention to the qualities of comprehensiveness, organization of project elements, and harmonious co-ordination between the elements. This demands reliance on logical criteria, by which these three principles are evaluated. The most general logical criteria remain as the most relevant for any field from which we may draw data as inputs into manifest variables that make up the latent variable model. Possibly the most general logical criteria remain as what have been described since the time of Aristotle as the laws of thought (Russell, 1912). These include three principles: the principle of identity, or something is what it is, the principle of non-contradiction (PNC) that no proposition can be both true and false at the same time, and the principle of the excluded middle, that a proposition is true, or false, and not anything else (PEM: Maddy, 2012). Findings within philosophy of mathematics suggest that identity and PNC remain as most relevant to all mathematical knowledge (see Maddy, 2012; Ferrerios, 2016). It is proposed here that the logical framework for research projects that include psychometric practices is well served by remaining simply reliant on propositions of identity, and propositions aimed at clarifying true/false criteria under PNC. An important aspect of utilizing identity and PNC in the logical framework that we are describing is that we are traversing links which cannot truly be said to make up a 'closed system' for that research project. The logic as used does not allow for the closure of logical models, or deductive statements to follow (see Suppes, 1999; Tarski, 1930). The project of creating a logical framework thus persists in a situation of

bounded rationality. Defined originally by Herbert Simon (1957), bounded rationality describes the constraints that apply to human decision making given limited information, time, and the limits of human cognition (Gigerenzer & Goldstein, 1996). Where studies demonstrate that simple heuristics can result in better quality decisions than those that result from time-consuming theoretically intensive approaches, bounded rationality invites the inclusion of systematized technique for risk calculations as relevant to distinct aspects of the research project. What this means is that even though deductive conclusions are not available for statements that result from the research project, the researcher can provide some estimate of potential risk involved with any conclusions from the project. Assessments of risk can be made for ideology, theory, models, variables, relations, data and phenomena, as well as the research situation. One can imagine that where risk estimates add up into a large overall coefficient for the project, the claims made from the project outcomes can be assessed in light of these risk estimates.

3.3 Axiomatic set theory

One worthwhile question to pursue is how to set out these elements of the project wherein identity and PNC can be asserted, and where risk can be assessed. Patrick Suppes has an extensive body of work culminating in his 2002 book *Representation and Invariance of Scientific Structures.* Early work in Suppes' corpus was aimed at clarifying the benefits of deductive closure within logical structures, particularly for physical sciences (see Suppes, 1951). By 2002 he advocated strongly for axiomatic practices that were informed by empirical practices as much as they were formal theory, working with a hierarchy of models as relevant to the relations between theory, and phenomena (Suppes, 1962; Boumans, 2016). This became advocacy for axiomatization to make explicit the structures of scientific theories and their models, and the real-world situations which ground their relevance (see Suppes, 2002). For Suppes, scientific practice is always aimed at tracking invariances in the phenomenal world, and axiom systems are proposed as the most effective and efficient way of setting clear the properties that structures must possess in the tracking of invariances. Intuitions about the existence of latent variables has given rise to an extraordinary array of model and software developments that facilitate the collection and analysis of statistical outcomes that may be pertinent to the existence of the phenomena and patterns that the latent variable notionally represents. Clarification via axiomatic structures gives us a means by which we can understand the constraints that should apply, in order that we know we are tracking real phenomena. In this approach the conceptual structure is made explicit by identifying the most basic elements for the given domain of the research project and writing axioms that express the essential properties for the elements, and the relations that hold, for the elements, within the domain. A second step, at the level of the framework, is to connect the elements across domains, where there is some relationship of import, and a third step is to assess any risk as appropriate for the elements, the domains, the relations, and the project as a whole. This information can be gathered in simple relational database form and made available in public domains, similar to the present-day preregistration processes under the Open Science Framework (Brandt et al., 2014).

3.4 Constructive mathematics

In *Defending the Axioms* (2011), Maddy makes a case for adopting a metaphysically-light or thin realism which attends to the constraints necessary for an axiom system that makes possible the

mathematical problem or solution under question, but no more. This is distinct from the metaphysically-weighted scientific realism that was proposed in the solution to the existence of latent variables for the latent variable model (Borsboom, Mellenbergh, & van Heerden, 2003; Borsboom, 2005; Hacking, 1983; Devitt, 1991). Maddy's (2011) approach is consistent with that of Ferrerios (2016), for whom the objectivity of mathematics is derived from a constructivist consideration of the kinds of constraints that give rise to the possibility of the relevance of the solutions. What this implies in connection to latent variable modelling in psychometric practices is that constraints set by each domain of the research project can be brought into a unified view. Claims for latent variables in light of problems such as the logical ones associated with factor indeterminacy can then be scrutinized according to the risk observed in the project framework. These constraints are articulated across world conditions, researcher stance, ideological commitments, scientific theory, scientific model, methodology, variables, relations, data and the phenomenon, itself, in relational database form.

3.5 Reproducibility questions

How does any of this help to address questions raised for psychometric practices connected to psychology research as raised in the outcomes of the Reproducibility Project (Open Science Collaboration, 2015)? In brief, the Reproducibility Project represented an attempt to replicate 100 studies first conducted across 2008-2009, across social and cognitive sciences. For these attempted replications "replication effects were half the magnitude of original effects, representing a substantial decline. … Thirty-six percent of replications had statistically significant results" (Open Science Collaboration, 2015, p. 4716–1). Subsequent analysis of the replication intention and efforts have suggested logical concerns to do with the expectations regarding replications themselves (Schmidt, 2010; Stanley & Spence, 2014). The logical framework described here may assist in answering logical questions for each project, such as ones regarding the nature of the psychological construct and its relation to psychological phenomena; or systematic aspects of measurement error which are otherwise not able to be tracked in either original studies, or attempted replications. This assessment may be used to feedback information regarding measurement error and/or construct to phenomenon relations, as raised in Schmidt (2010) and Stanley and Spence (2014).

3.6 Rough set theory

Set theory offers other resources for gathering non-statistically based-evidence for aspects of psychological phenomena. Rough set theory was devised by Zdzisław Pawlak (1982) as a method that does not rely on anything other than the data itself to track patterns or generate decision rules. Two specific applications of rough set theory as preliminary steps to statistical analyses in latent variable modelling may be i) generating an analysis for decisions regarding the relevance of variables for any possible model; ii) digging into data otherwise classed as measurement error to track otherwise indiscernible patterns in those data. Information regarding the relevant variables for an analysis typically otherwise only comes from substantive theory (MacCallum & Austin, 2000). An algorithm-based analysis that tracks key relations between variables may be helpful in asserting that relevant variables have been included in the models referenced by the research project. The second use for latent variable theory involves analysis of what is otherwise classed as measurement error in a latent variable model.

This may be relevant for example in Rasch latent variable modelling, where linearity in relations is assumed between items on a test. It may be the case that there are patterns of non-response or no response in the data that are elided in the overall Rasch analysis which would otherwise not come to the attention of the researcher (see Andersen, 1995). Utilising user-specified upper and lower approximations of what should be included in a set, groupings or subsets that exhibit non-linear patterns amongst item responses can be selected from datasets, and further analysed by researchers. This would take place in the final steps of the research project but may serve as a useful adjunct to reporting of statistical outcomes where otherwise these patterns would be subsumed into measurement error.

4 Conclusions

The logical framework as presented here describes the structure of a theory of expected relations for any research project that makes use of the latent variable model in psychometric analysis. The latent variable model has been demonstrated as extending to most statistical modelling techniques utilized in present day psychometrics (McDonald, 1999; Skrondal & Rabe-Hesketh, 2004; Mulaik, 2010). Techniques that do not fall under the latent variable model may still place much weight on an assumption of conditional independence, where for example in dynamic network modelling the conditional independence assumption is shifted to every possible pairwise relationship between any possible variables included in the model (Epskamp, Borsboom, & Fried, 2018). Any reliance on such an assumption is well served by non-statistical evidence and evidentiary methods that support identification of the phenomena by elaborating the logical constraints most relevant to the existence of the phenomena. Where mathematics makes available to all sciences the most certain of knowledge, as well as perhaps the most abstract of techniques, we would be well-served in our psychometric practices if we choose to be informed by the best of what mathematics stands to teach us. Here, ontology may be used as a test ground for the existence of phenomena and not as a declaration ground for existence under practice, by simple fiat.

Acknowledgments

References

Andersen, E. B. (1995). What Georg Rasch would have thought about this book. In G.H Fiscer & I.W. Molenaar (Eds.). Rasch Models (pp. 383-390). New York: Springer-Verlag.

Bartholomew, D. J. (1987). Latent Variable Models and Factor Analysis. London, UK:Griffin.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. Sociological Methods & Research, 16(1), 78-117. DOI: doi/abs/10.1177/0049124187016001004

Boag, S. (2011). Explanation in personality psychology:"Verbal magic" and the five-factor model. Philosophical Psychology, 24(2), 223-243.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. Annual Review of Psychology,53, 605-634. doi: 10.1146/annurev.psych.53.100901.135239

Borsboom, D. (2005). Measuring the Mind: Conceptual Issues in Contemporary Psychometrics. Cambridge: Cambridge University Press.

Borsboom, D. (2006). The attack of the psychometricians. Psychometrika, 71, 425-440. doi: 10.1007/s11336-006-1447-6

Borsboom, D. (2008). Latent variable theory. Measurement: Interdisciplinary Research and Perspectives, 6, 25-53. doi: 10.1080/15366360802035497

Borsboom, D., & Mellenbergh, G. J. (2004). Why psychometrics is not pathological: A comment on Michell. Theory & Psychology, 14(1), 105-120.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. Psychological Review, 110, 203-219.

Boumans, M. (2016). Suppes's outlines of an empirical measurement theory, Journal of Economic Methodology, 23:3, 305-315, DOI: 10.1080/1350178X.2016.1189124

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., ... & Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication?. Journal of Experimental Social Psychology, 50, 217-224. DOI: https://doi.org/10.1016/j.jesp.2013.10.005

Cliff, N. (1983). Some Cautions Concerning The Application Of Causal Modeling Methods, Multivariate Behavioral Research, 18, 115-126, DOI: 10.1207/s15327906mbr1801_7

Costa, P. T., & McCrae, R. R. (2008). The revised neo personality inventory (NEO-PI-R). The SAGE Handbook of Personality Theory and Assessment, 2, 179-198.

Cowles, M. (2001). Statistics in Psychology: An Historical Perspective. London: Psychology Press.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. Behavior Research Methods, 50(1), 195-212.

Eusebi, P., Reitsma, J. B., & Vermunt, J. K. (2014). Latent class bivariate model for the meta-analysis of diagnostic test accuracy studies. BMC Medical Research Methodology, 14, 88. doi:10.1186/1471-2288-14-88

Fayers, P. M., & Hand, D. J. (1997). Factor analysis, causal indicators and quality of life. Quality of Life Research, 6, 139-150. doi: 10.1023/A:1026490117121

Ferreirós, J. (2016). Mathematical Knowledge and the Interplay of Practices. Princeton: Princeton University Press

Galton, F. (1879). Psychometric experiments. Brain: A Journal of Neurology, 11, 149-162.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. Psychological review, 103(4), 650-669.

Guttman, L. (1950). The Basis for Scalogram Analysis. In S. A. Stouffer et al. (Eds). Measurement and Prediction. Studies in Social Psychology in World War II (pp. 60- 90). Princeton, NJ: Princeton University Press.

Guttman, L. (1955). The determinacy of factor score matrices with implications for five other basic problems of common-factor theory. British Journal of Statistical Psychology, 8, 65-81.

Hacking, I. (1983). Representing and Intervening (Vol. 279). Cambridge: Cambridge University Press.

Haig, B. D. (2009). Inference to the best explanation: A neglected approach to theory appraisal in psychology. The American journal of psychology, 219-234.

Harman, G. H. (1965). The inference to the best explanation. The philosophical review, 74(1), 88-95.

Hempel, C.G. (1965). Aspects of Scientific Explanation and other Essays in the Philosophy of Science. New York: Free Press.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. Psychometrika, 36, 109–133.

Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O.D. Duncan (Eds.), Structural Equation Models in the Social Sciences (pp. 85-112). New York: Seminar Press.

Kaufman, A. S. (1979). Intelligent Testing. WISC-R. New York: Wiley-Interscience.

Lord, F. M., & Novick, M. R. (1968). Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley.

Luce, R. D. (1997). Quantification and symmetry: Commentary on Michell, Quantitative science and the definition of measurement in psychology. British Journal of Psychology, 88(3), 395-398.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. Annual review of psychology, 51(1), 201-226. DOI: 10.1146/annurev.psych.51.1.201

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. Psychological Bulletin, 114(1), 185-199. DOI: http://dx.doi.org/10.1037/0033-2909.114.1.185

Maddy, P. (2011). Defending the Axioms: On the Philosophical Foundations of Set Theory. Oxford: Oxford University Press.

Maddy, P. (2012). The philosophy of logic. Bulletin of Symbolic Logic, 18(4), 481-504. DOI: https://doi.org/10.2178/bsl.1804010

Maraun, M. D. (1996). Metaphor taken as math: indeterminancy in the factor analysis model. Multivariate Behavioral Research, 31(4), 517-538. doi: 10.1207/s15327906mbr3104_6

Maraun, M. D., & Gabriel, S. M. (2013). Illegitimate concept equating in the partial fusion of construct validation theory and latent variable modeling. New Ideas in Psychology, 31, 32-42. doi: 10.1016/j.newideapsych.2011.02.006

McDonald, R.P. (1999). Test Theory. Mahwah: NJ: Lawrence Erlbaum Associates.

McDonald, R. P. (2003). Behavior domains in theory and in practice [Measurement for the Social Sciences: Classical Insights into Modern Approaches]. Alberta Journal of Educational Research, 49(3), 212-230.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. Psychometrika, 58(4), 525-543. DOI: 10.1007/BF02294825

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. British Journal of Psychology, 88(3), 355-383.

Michell, J. (2000). Normal science, pathological science and psychometrics. Theory & Psychology, 10(5), 639-667. DOI: doi/abs/10.1177/0959354300105004

Michell, J. (2003). The quantitative imperative: Positivism, naïve realism and the place of qualitative methods in psychology. Theory & Psychology, 13(1), 5-31.

Mulaik, S. A. (2010). Foundations of Factor Analysis (2nd Edition). Boca Raton, Florida: Chapman & Hall/CRC.

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. Perspectives on Psychological Science, 7(6), 657-660.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. Science, 349(6251), aac4716.

Pawlak, Z. (1982). Rough sets. International journal of computer & information sciences, 11(5), 341-356.

Petocz, A., & Newbery, G. (2010). On conceptual analysis as the primary qualitative approach to statistics education research in psychology. Statistics Education Research Journal, 9, 123-145.

Rescher, N. (2016). Inference from the best systematization. Mind & Society, 15(2), 147-154.

Romeijn, J. W., & Williamson, J. (2018). Intervention and Identifiability in Latent Variable Modelling. Minds and Machines, 1-22. DOI: 10.1007/s11023-018-9460-y

Salmon, W. C. (1990). Four Decades of Scientific Explanation. Minnesota: Regents of the University of Minnesota.

Schmidt, F. (2010). Detecting and correcting the lies that data tell. Perspectives on Psychological Science, 5(3), 233-242.

Sijtsma, K. (2006). Psychometrics in psychological research: Role model or partner in science?. Psychometrika, 71(3), 451-455. DOI: https://doi.org/10.1007/s11336-006-1497-9

Simon, H.A. (1957. Models of Man. New York: Wiley.

Skrondal, A. & Rabe-Hesketh, S. (2004). Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Modelling. Boca Raton, FL: Chapman & Hall/CRC.

Spearman, C. (1904). " General Intelligence," objectively determined and measured. The American Journal of Psychology, 15(2), 201-292.

Spearman, C. (1929). The uniqueness of "g". Journal of Educational Psychology, 20(3), 212-216. http://dx.doi.org/10.1037/h0072998

Stanley, D. J., & Spence, J. R. (2014). Expectations for replications: Are yours realistic?. Perspectives on Psychological Science, 9(3), 305-318.

Stigler, S. M. (2007). The epic story of maximum likelihood. Statistical Science, 22, 598-620.

Suppes, P. (1951). A set of independent axioms for extensive quantities. Portugaliae Mathematica, 10 (4), 163-172.

Suppes P. (1962). Models of data. In E. Nagel, P. Suppes, A. Tarski (Eds.). Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress. Amsterdam: Stanford University Press, pp. 252–261.

Suppes, P. (1999). Introduction to Logic. New York: Dover Publications.

Suppes, P. (2002). Representation and Invariance of Scientific Structures. Stanford: CSLI publications.

Suppes, P. (2006). Transitive indistinguishability and approximate measurement with standard finite ratio-scale representations. Journal of Mathematical Psychology, 50(3), 329-336.

Suppes, P. (2016). Qualitative axioms of uncertainty as a foundation for probability and decision-making. Minds and Machines, 26(1-2), 185-202.

Tarski, A. (1930). Fundamentale Begrie der Methodologie der deduktiven Wissenschaften. I. Monatshefte für Mathematik und Physik, 37, 361-404.

Thomson, G. H. (1916). A hierarchy without a general factor. British Journal of Psychology, 8, 271-281.

Thorndike, E. L. (1927). Measurement of Intelligence. New York: Bureau of Publications, Teachers College, Columbia University.

Thurstone, L. L. (1934). The vectors of mind. Psychological review, 41(1), 1-32. DOI: http://dx.doi.org/10.1037/h0075959

Weinberger, N. (2015). If intelligence is a cause, it is a within-subjects cause. Theory & Psychology, 25(3), 346-361. DOI: doi/abs/10.1177/0959354315569832

Wilson, E B. (1928). Review of 'The Abilities of Man, Their Nature and Measurement' by C. Spearman. Science, 67, 244-248.