The science in human science research: The case for Rasch measurement in learning environment research

R F Cavanagh1 and W P Fisher Jr2
1 School of Education, Curtin University, GPO Box U1987, Perth, Western Australia
2 BEAR Center, Graduate School of Education, University of California, Berkeley, Berkeley California 94702 USA

Email: R.Cavanagh@curtin.edu.au

Abstract. The field of learning environment research has a 40 year history originating in the USA with the work of Herbert Walberg and Rudolf Moos, who utilised Kurt Lewin’s Field Theory. The conception of a learning environment employed in learning environment research has social and psychological aspects with a focus on learning, student achievement, and student attitudes. In 1984, the American Educational Research Association (AERA) special interest group (SIG) on the Study of Learning Environments was formed, and in 1998 the first edition of the Learning Environments Research journal (LERj) was published. This paper examines the features of the survey instruments developed to measure aspects of classrooms and other environments where learning occurs. The traditions of Classical Test Theory, True-Score Theory and Rasch's models for unidimensional measurement are related to learning environment research as reported in the LERj. Measurement theories are associated with three philosophical orientations: positivism, anti-positivism, and post-positivism. The paper concludes with some examples of how a measurement science extending Rasch's models can be applied in learning environment research.

1. History of learning environment research

The genesis of the field of learning environments can be traced back four decades to the pioneering work of Herbert Walberg and Rudolf Moos. The Harvard Project Physics evaluation stimulated Walberg to develop the Learning Environment Inventory [1]. Concurrently, Moos [2,3] developed social climate scales to profile human environments. The underlying theory for both lines of inquiry was Field Theory, which focuses on human behaviour resulting from the interaction of personal characteristics and the environment [4]. Applying this notion to learning required defining the specifics of the environment in which it takes place. Fraser [5 p. 3] viewed the learning environment as “the social, psychological and pedagogical contexts in which learning occurs and which affect student achievement and attitudes”. It encompasses formal educational settings: kindergarten, primary, secondary and tertiary schooling; within school and class programs; and on-line instruction. It also currently includes less formal settings: the home, school playground, sport, computer gaming, internet activities, and extra-curricular activities.

Numerous applications of learning environment assessments are found in school psychology; in evaluations of educational programs, innovations and provisions; in associations between the classroom environment and students’ cognitive and effective outcomes [6,7]; and as feedback in teachers’ action research attempts to change their classrooms [8].

A significant event in the history of learning environment research occurred in 1984 when the American Educational Research Association (AERA) special interest group (SIG) on the Study of Learning Environments was constituted. The SIG has provided a vehicle for the dissemination of...
2. Quantification in learning environments research

The scope of settings and of the participants in learning environment research is extensive; however, there is a clear overarching intent. The research centres on investigating associations between students’ outcomes and their perceptions of classroom environments [5]. Methodologically, early research applied survey methods with instrument scales based on a priori construct models. In time, these methods were supplemented by a more interpretative exploratory approach. However, survey instruments are still extensively used and some of the earlier instruments are still in use.

Learning environment instruments typically have the following features: Moos’ 1974 schema; a classroom learning focus; eliciting student perceptions; seeking student affirmation of item statements; rating scale response categories (assumed ordinal); actual and preferred scales; some negatively worded items; raw data aggregated at the scale level; instrument construction using Exploratory Factor Analysis; Coefficient Alpha reliability estimation; inter-scale correlations interpreted as indicating discriminant validity; and ANOVA applied to test for group (e.g. cross-cultural) differences [5]. The multi-variate nature of the data has led to application of Multiple Analysis of Variance (MANOVA), Confirmatory Factor Analysis (CFA), and Structural Equation Modelling (SEM). Additionally, research into environmental influences accounting for classroom level learning effects, such as teacher attributes and school-level factors, employs Hierarchical Linear Modelling (HLM). These analytic techniques share a common theoretical origin, the Test-score Tradition of Measurement (Classical Test Theory and True-score Theory), which “focuses on test scores or summed item responses with the main theme of quantifying measurement error and sources of error variance in measurement” [9, p. 104].

While the dominant measurement approach in learning environment research is the Test-score Tradition, LERj also includes examples of the Scaling Tradition of Measurement, typically involving Rasch models [10]. “The scaling tradition focuses on the calibration of individual items and the measurement of persons on a shared latent variable” [11, p. 126]. The first Scaling Tradition learning environment publication in LERj was a 2002 study of parent receptivity [12]. This was one of only eleven Rasch measurement articles published in the LERj between 1998 and 2017. In comparison, over the same period, 208 articles using factor analytic methods were published. According to LERj reports, Rasch measurement has been employed to: calibrate item difficulty measures and person ability measures on the same linear scale; produce item difficulty estimates independent of the sample of respondents; produce person ability estimates independent of the items administered; order items from ‘easy’ to ‘hard’ and persons from low ability to high ability; test whether all the items measure the same latent construct; test for logical use of rating scale categories; score dichotomous and polytomous data on the same linear scale; produce graphical displays; and test construct models.

3. A philosophical perspective on learning environment measurement

Examination of particular measurement procedures and principles can indicate the philosophical origins of a research methodology. Associations between learning environment research methods and three philosophical orientations—positivism, anti-positivism, and post-positivism—can be discerned in the overview of the field. Reconciling conceptions of science, theories of measurement, and research methods, requires articulate contrasts of how differing "philosophies of science inform paradigmatically distinct measurement principles and procedures” [13]. Situating current methods in learning environments research relative to one another in the context of turns in the history and philosophy of science aids in grounding our understandings of representations of science across fields.
First, positivism: positivism assumes a universal and a priori scientific method; an objective, independent reality called the world; an overall discernible order in the world; truth as correspondence to reality; a truth discovered by scientists functioning as spectators of the given world; reductionism, in principle, as possible and desirable; facts as unambiguously distinguished from values; and theoretical statements translatable into statements about observation (observation/theory distinction) [14]. The assumptions incumbent in positivism can be identified in Test-score Theory in general [13], and specifically in the methods applied in learning environment Test-score Theory research. For example, multi-variate data analysis using Factor Analysis and Structural Equation Modelling apply correlational methods (including modelling covariance) to deductively confirm postulated relations between variables. In these methods, nonlinear ordinal data are typically assumed to be interval measures in the absence of any tests of a defined unit, and the meaning of the measured construct is left unarticulated. A positivist view of science is consistent with the application of Test-score Theory in learning environment research.

Second, anti-positivism: Anti-positivism strongly rejects positivist principles [15,16]. Some of the criticisms include rejection of differentiation between subjects and objects, mind and world, language and reality as possible and desirable; embracing inter-subjective meanings as embodying the discourse of personal agency, moral obligation and political responsibility; accepting scientific methods and practices as local and not universal; and accepting that observations cannot be completely independent of theory. During the 1990s, learning environment research was influenced by the anti-positivist movement as reflected in the growth of naturalistic approaches. However, there was pragmatic coexistence between the traditional survey approach and the more recent interpretive approach; the positivist approach escaped serious scrutiny. The anti-positivist view of science does not appear to have had a major influence on measurement in learning environment research.

Third, post-positivism: Post-positivism is often conceptualised and justified as a response to the anti-positivism movement, as a positive reaction to the deficit nature of anti-positivism and the criticism of positivism. It is also the dominant paradigm in much human and social science research. However, in learning environment measurement, the strong alignment of Test-score Theory learning environment measurement with positivism may limit the application of post-positivist principles and methods. One way to overcome this limitation is to articulate the ways in which Rasch measurement science can be made consistent with a post-positivist view of science [17]. This approach has four elements: it advocates measurement that enables societal and environmental renewal; a philosophical orientation of hermeneutic phenomenology; adopting constructs from related disciplines, including network theory, systems theory and metrology; and applies Rasch measurement science not as a primarily data-analytic method, but as oriented to construct theory, rich interpretive principles, incorporation of natural language's levels of complexity, and the distribution to end users of instruments calibrated to be traceable to defined units of comparison.

4. Applying Rasch measurement science in learning environment research

Applying Rasch measurement science has the potential to improve learning environment research. For example: Persons and items (representing the learning environment) are not separated, as they are in Field Theory; characteristics of persons and of the environment are qualified and quantified; the scales are invariant, not person or item dependent; the technical limitations of Classical Test Theory and True-score Theory are avoided; construct models are stochastic and not deterministic; philosophical and theoretical foundations for procedures are specified; and learning progressions can be modelled. Most rating scale numbers subjected to statistical comparisons in learning environments research have not been evaluated to see (a) if they express measures of a single unidimensional construct in a linear, additive unit, and (b) if they exhibit the properties of precision and meaningfulness needed to make them fit for purpose in applications. The advantages of linearizing rating scales' ordinal units, as shown in Figure 1, include being able to take missing data into account, adaptively administering instruments in tailored applications, equating different instruments to measure in a common unit, the development of parallel forms or item banks, etc. [18]. These advantages comprise a feasible, viable, and desirable
research agenda for the field of learning environments research, and will be pursued in forthcoming investigations.

![Figure 1. Ordinal scores and interval measures](image)

5. References

[1] Walberg H J 1979 *Educational Environments and Effects* (San Pablo, CA: McCutchan)

[2] Moos R W 1974 *The Social Climate Scales* (Palo Alto, CA: Consulting Psychologists Press)

[3] Moos R W 1979 *Evaluating Educational Environments* (San Francisco, CA: Jossey Bass)

[4] Lewin K 1939 *Am J Soci.* 44 868-896

[5] Fraser B J 1998 *Learn Environ Res* 1 1-5

[6] Fraser B J 2012 Classroom learning environments: Retrospect, context and prospect In B J Fraser, K G Tobin and C J McRobbie (Eds) *Second international handbook of science education* pp. 1191-1232 (New York: Springer.)

[7] Fraser B J 2014 Classroom learning environments: Historical and contemporary perspectives In N G Lederman and S K Abell (Eds), *Handbook of research on science education, Volume II* pp. 104-119 (New York: Routledge)

[8] Fraser B J and Aldridge J M 2017 Improving classrooms through assessment of learning environments In J P Bakken (Ed) *Classrooms Volume 1: Assessment practices for teachers and student improvement strategies* pp. 91-107 (New York: Nova)

[9] Engelhard G Jr 2012 *Invariant Measurement* (New York: Routledge Academic)

[10] Rasch G 1960 *Probabilistic Models* (Copenhagen, Denmark: Danmarks Paedagogiske Institut)

[11] Engelhard G Jr 1991 *Rasch Meas Trans* 4 12

[12] Waugh R F and Cavanagh R F 2002 *Learn Environ Res* 5 329-352

[13] Cavanagh R F and Fisher W P Jr In-press 2018 Research design in human science research *Proceeding of the Pacific Rim Objective Measurement Symposium 2017* (Berlin Heidelberg: Springer-Verlag)

[14] Peters M A and Burbules N C 2004 *Poststructuralism and Educational Research* (Lanham, MD: Rowman & Littlefield Publishers, Inc)

[15] Kuhn T 1970 *The Structure of Scientific Revolutions.* (Chicago: University of Chicago Press)

[16] Quine W V O 1951 *Philos Rev* 60 20-43

[17] Fisher W P Jr and Cavanagh R F 2016 Measurement as a medium for communication and social action I & II *Proceedings of the Pacific Rim Objective Measurement Symposium 2015* (Berlin Heidelberg: Springer-Verlag)

[18] Wright B D and Stone M H 1999 *Measurement Essentials* (Wilmington, DE: Wide Range, Inc)