

Measurement Science in Psychology: Prospects for New SI Units

William P. Fisher, Jr.

Could the instruments of psychology and the social sciences be tuned to common scales? How might humanity's relationships with the earth be harmonized for diverse ensembles? Can different tuning systems inform different cultural perspectives? In what ways might people, communities, and nations make beautiful music together, creatively improvise bent and blue notes played out in jazzy, swinging, and rocking social, economic, and political organizations?

L. L. Thurstone was a former electrical engineer turned psychologist who made foundational contributions to measurement theory and practice. In an article published in 1928, Thurstone wrote:

"One crucial experimental test must be applied to our method of measuring attitudes before it can be accepted as valid. A measuring instrument must not be seriously affected in its measuring function by the object of measurement. To the extent that its measuring function is so affected, the validity of the instrument is impaired or limited. If a yardstick measured differently because of the fact that it was a rug, a picture, or a piece of paper that was being measured, then to that extent the trustworthiness of that yardstick as a measuring device would be impaired. Within the range of objects for which the measuring instrument is intended, its function must be independent of the object of measurement."

And so, Thurstone held that, in psychological measurement:

"If the scale is to be regarded as valid, the scale values of the statements should not be affected by the opinions of the people who help to construct it. This may turn out to be a severe test in practice, but the scaling method must stand such a test before it can be accepted as being more than a description of the people who construct the scale."

This article has been cited in peer-reviewed research over 3,200 times. Thurstone's description of a principle of substantive unit amounts that add up in a way usefully represented by numbers set the stage for many later developments in psychological and social measurement. Rasch (1960, 1961), in particular, formalized Thurstone's ideas in individual-level models based on a parameter separation theorem, where the observed score is both necessary and sufficient to estimation (Andersen, 1977; Andrich, 1978, 2010; Fischer, 1981).

Rasch's work has formed the basis of recent collaborations of metrologists and psychometricians spelling out the terms of how measurement models, methods, concepts, and unit standards could be unified across the sciences (Cano, et al., 2019; Fisher & Cano, 2023; Mari & Wilson, 2014; Mari, Wilson, & Maul, 2021; Pendrill, 2014, 2019; Pendrill & Fisher, 2015).

Rasch's models for measurement are notable for:

- extending the implicit mathematics of everyday language into explicitly mathematical language in much the same way this was accomplished in the natural sciences;
- having the same mathematical form as many laws of nature, including intriguing implementations of multivalued quantum logic and nonequilibrium evolutionary processes;
- reproducing physical measurements of mass, length, and density from ordinal observations;
- estimating the same quantities as metrological methods when applied to the same data;
- being recognized by metrologists as paradigmatic of measurement;
- integrating explanatory models' theoretical predictions with experimental tests of units defined as retaining their properties across samples and instrument configurations;
- being applied in tens of thousands of published research articles;

- ensuring the defensibility of hundreds of millions of admissions, graduation, licensure, and certification decisions over the last 50 years and more;
- humanizing and personalizing quantitative psychology and social science by relating measurement to the fulfillment of educational, career, and health goals; and
- supporting, perhaps unexpectedly to most, a human, socially progressive, aesthetic, ethical, economical, and environmentally sustainable epistemology of science.

Tuning the instruments of psychology and the social sciences to common scales offers hopeful possibilities for harmonizing relationships in creative ways that do not reduce beauty and meaning to homogenized uniformity but instead open onto new playful improvisations and soulful innovations.

These points will be briefly elaborated in a presentation, and will be supported with referenced sources for those wishing to pursue their own interests in the relevant models and methods.

References

- Andersen, E. B. (1977). Sufficient statistics and latent trait models. *Psychometrika*, 42(1), 69-81.
- Andrich, D. (1978, December). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573.
- Andrich, D. (1978). Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement*, 2, 449-460.
- Andrich, D. (2010). Sufficiency and conditional estimation of person parameters in the polytomous Rasch model. *Psychometrika*, 75(2), 292-308.
- Cano, S., Pendrill, L., Melin, J., & Fisher, W. P., Jr. (2019). Towards consensus measurement standards for patient-centered outcomes. *Measurement*, 141, 62-69.
- Fischer, G. H. (1981, March). On the existence and uniqueness of maximum-likelihood estimates in the Rasch model. *Psychometrika*, 46(1), 59-77.
- Fisher, W. P., Jr., & Cano, S. (Eds.). (2023). *Person-centred outcome metrology: Principles and applications for high stakes decision making* (G. R. R. Morawski, & others, Ed.). Springer Series in Measurement Science & Technology. Springer. <https://link.springer.com/book/9783031074646>
- Mari, L., & Wilson, M. (2014, May). An introduction to the Rasch measurement approach for metrologists. *Measurement*, 51, 315-327.
- Mari, L., Wilson, M., & Maul, A. (2021). *Measurement across the sciences: Developing a shared concept system for measurement* (R. Morawski, G. Rossi, & others, Eds.). Springer Series in Measurement Science and Technology. Springer.
- Pendrill, L. R. (2014, December). Man as a measurement instrument [Special Feature]. *NCSLi Measure: The Journal of Measurement Science*, 9(4), 22-33.
- Pendrill, L. R. (2019). *Quality assured measurement: Unification across social and physical sciences*. Springer.
- Pendrill, L., & Fisher, W. P., Jr. (2015). Counting and quantification: Comparing psychometric and metrological perspectives on visual perceptions of number. *Measurement*, 71, 46-55.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests* (Reprint, with Foreword and Afterword by B. D. Wright, Chicago: University of Chicago Press, 1980). Danmarks Paedagogiske Institut.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In J. Neyman (Ed.), *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability: Volume IV: Contributions to biology and problems of medicine* (pp. 321-333 [<http://www.rasch.org/memo1960.pdf>]). University of California Press.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, XXXIII, 529-544. (Rpt. in L. L. Thurstone, (1959). *The measurement of values* (pp. 215-233). Chicago, Illinois: University of Chicago Press, Midway Reprint Series.)