

Psychometrics and Measurement

- Kane, M.T. (2013), Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*, 50: 1-73.
<https://doi.org/10.1111/jedm.12000>
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological review*, 111(4), 1061–1071.
<https://doi.org/10.1037/0033-295X.111.4.1061>
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, 71(3), 425–440. <https://doi.org/10.1007/s11336-006-1447-6>
- Loevinger, J. (1957). Objective Tests as Instruments of Psychological Theory. *Psychological Reports*, 3(3), 635–694. <https://doi.org/10.2466/pr0.1957.3.3.635>
- Bringmann, L. F., & Eronen, M. I. (2016). Heating up the measurement debate: What psychologists can learn from the history of physics. *Theory & Psychology*, 26(1), 27–43. <https://doi.org/10.1177/0959354315617253>
- Trendler, G. (2009). Measurement Theory, Psychology and the Revolution That Cannot Happen. *Theory & Psychology*, 19(5), 579–599.
<https://doi.org/10.1177/0959354309341926>
- Vessonen, E. (2019). Operationalism and realism in psychometrics. *Philosophy Compass* 14 (10).
- Sijtsma, K. (2012). Psychological measurement between physics and statistics. *Theory & Psychology*, 22(6), 786–809.
<https://doi.org/10.1177/0959354312454353>
- Michell, J. (2021). Representational measurement theory: Is its number up? *Theory & Psychology*, 31(1), 3–23. <https://doi.org/10.1177/0959354320930817>
- Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110, 305-314.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110, 203-219
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Michell, J. (2008). Is psychometrics pathological science? *Measurement*, 6, 7-24.
- Morris, S. D., Grice, J. W., & Cox, R. A. (2016). Scale imposition as quantitative alchemy: Studies on the transitivity of neuroticism ratings. *Basic and Applied Social Psychology*.
- van Bork, R., Rhemtulla, M., Sijtsma, K., & Borsboom, D. (preprint). A causal theory of error scores. doi:10.31234/osf.io/h35sa (preprint: <https://psyarxiv.com/h35sa/>)

- Jonas, K., & Markon, K. E. (2016). A descriptivist approach to trait conceptualization and inference. *Psychological Review*, 123 (1) 90-96
- Lord, F. M. (1953). On the statistical treatment of football numbers. *American Psychologist*, 8(12), 750–751. <https://doi.org/10.1037/h0063675>
- Maul, A. (2017). Rethinking traditional methods of survey validation. *Measurement: Interdisciplinary Research and Perspectives*, 15(2), 51-69.
- Scholten, A. Z., & Borsboom, D. (2009). A reanalysis of Lord's statistical treatment of football numbers. *Journal of Mathematical Psychology*, 53(2), 69-75.
- Van der Maas, H. L., Kan, K. J., & Borsboom, D. (2014). Intelligence is what the intelligence test measures. *Seriously. Journal of Intelligence*, 2(1), 12-15.
- Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Voodoo correlations in social neuroscience. *Perspectives on Psychological Science*, 4(3), 274-290.
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1(2), 199.
- Schimmack, U. (2021). The validation crisis in psychology. *Meta-Psychology*, 5.
- Borsboom, D., Cramer, A. O., Kievit, R. A., Scholten, A. Z., & Frančić, S. (2009). The end of construct validity. In *The Concept of Validity: Revisions, New Directions and Applications*, Oct, 2008. IAP Information Age Publishing.
- Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456-465.
- Brown, G.. (2021). Principles and assumptions of psychometric measurement (Version 1). The University of Auckland. <https://doi.org/10.17608/k6.auckland.17057828.v1>
- Tryon, R. C. (1935). A theory of psychological components—an alternative to "mathematical factors." *Psychological Review*, 42(5), 425–454. <https://doi.org/10.1037/h0058874>
- MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, 55(2), 95–107. <https://doi.org/10.1037/h0056029>
- Stanford Encyclopedia of Philosophy - Measurement in Science <https://plato.stanford.edu/entries/measurement-science/>
- Suppes, P., & Zinnes, J. L. (1962). Basic Measurement Theory. Technical Report No. 45. (March 15, 1962). https://web.stanford.edu/group/csli-suppes/techreports/IMSSS_45.pdf

- Rozeboom, W. W. (1966). Scaling theory and the nature of measurement. *Synthese*, 170-233.
<https://link.springer.com/content/pdf/10.1007/BF00485356.pdf>
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of mathematical psychology*, 1(1), 1-27. [https://doi.org/10.1016/0022-2496\(64\)90015-X](https://doi.org/10.1016/0022-2496(64)90015-X)
- Narens, L., & Luce, R. D. (1993). Further comments on the “nonrevolution” arising from axiomatic measurement theory. *Psychological Science*, 4(2), 127-130. <https://doi.org/10.1111/j.1467-9280.1993.tb00475.x>
- Narens, L., & Luce, R. D. (1986). Measurement: The theory of numerical assignments. *Psychological Bulletin*, 99(2), 166–180.
<https://doi.org/10.1037/0033-2909.99.2.166>
- Luce, R. D. (1996). The ongoing dialog between empirical science and measurement theory. *Journal of mathematical psychology*, 40(1), 78-98.
<https://doi.org/10.1006/jmps.1996.0005>
- "Luce, R. D., & Suppes, P. (2002). Representational measurement theory. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology: Methodology in experimental psychology* (pp. 1–41). John Wiley & Sons Inc.
<https://doi.org/10.1002/0471214426.pas0401>
- Luce, R. D., & Kruschke, J. K. (1988). Measurement, scaling, and psychophysics. *Stevens' handbook of experimental psychology*, 1, 3-74.
http://music.psych.cornell.edu/articles/math/Luce-Kruschke_Book%20Chapter_1988.pdf
- Engelhard Jr, G. (2008). Historical perspectives on invariant measurement: Guttman, Rasch, and Mokken. *Measurement*, 6(3), 155-189.
<https://doi.org/10.1080/15366360802197792>
- Proulx, T., & Morey, R. D. (2021). Beyond Statistical Ritual: Theory in *Psychological Science*. *Perspectives on Psychological Science*, 16(4), 671–681.
<https://doi.org/10.1177/17456916211017098>
- Luce, R. D., & Narens, L. (1987). Measurement scales on the continuum. *Science*, 236(4808), 1527-1532. <https://doi.org/10.1126/science.236.4808.1527>
- Steele, J. (2023, January 22). Can we measure effort in cognitive tasks? Examining the application of Additive Conjoint Measurement and the Rasch model. <https://doi.org/10.31234/osf.io/6pvht>
- Roberts, F. S. (1985). Applications of the theory of meaningfulness to psychology. *Journal of Mathematical Psychology*, 29(3), 311-332.
[https://doi.org/10.1016/0022-2496\(85\)90011-2](https://doi.org/10.1016/0022-2496(85)90011-2)
- Wright GJ, Feinstein AR. A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol*. 1991;45 (11):1201-1218. 10.1016/0895-4356(92)90161-f

- Haig, B. (2022). Repositioning construct validity theory: From nomological networks to pragmatic theories, and their evaluation by explanatory means. PsyArXiv. <https://doi.org/10.31234/osf.io/k54b6>
- Student, S. R., & Read, W. S. (2025). Applying Bayesian checks of cancellation axioms for interval scaling in limited samples. *Behavior Research Methods*, 57(11), 305. <https://doi.org/10.3758/s13428-025-02844-7>
- Scheiblechner, H. (1995). Isotonic ordinal probabilistic models (ISOP). *Psychometrika*, 60(2), 281-304. <https://doi.org/10.1007/BF02301417>
- Scheiblechner, H. (1999). Additive conjoint isotonic probabilistic models (ADISOP). *Psychometrika*, 64(3), 295-316. <https://doi.org/10.1007/BF02294297>
- Scheiblechner, H. (2007). A unified nonparametric IRT model for d-dimensional psychological test data (d-ISOP). *Psychometrika*, 72(1), 43-67. <https://doi.org/10.1007/s11336-005-1282-1>