Thurstone:
Measurement For a New Science

Nikolaus Bezruczko, Ph.D.

"He stole fire from the gods, then paid with factor analysis."

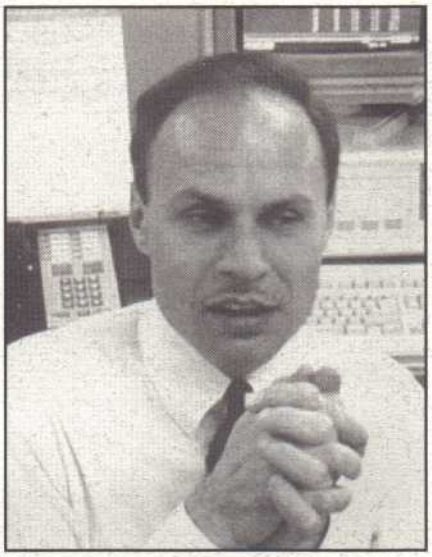
In ten short years, early in his career, Louis L. Thurstone revolutionized nonphysical scaling by single-handedly adapting the psychophysics developed by Fechner, Wundt, and Müller to measure mental forces in 20th century psychology. In contrast, the long, slow labor of factor analysis overwhelmed him for more than twenty years, as he tried to develop and defend it. His measurement advances were spectacularly laying the foundations for modern psychometrics, while factor analysis was a dismal burden, consuming his energy and distracting his attention. Scholars may argue whether factor analysis wasted his time, but all agree he never returned to absolute scaling.

Thurstone's contributions to social science, however, go deeper than inventing modern psychometrics. His goal was an entirely new theoretical psychology based on instincts, needs, and aspirations "where the dynamic self finds overt expression" (1923, 356), "We should analyze . . . [human actions] . . . as the expression of cravings that originate in the organism and find particular modes of satisfaction in the stimuli that happen to be available" (1923, 368). In Thurstone's brave new cosmology, psychology studies the objective representation of these mental forces, his alternative to stimulus-response behaviorism and subconscious psychoanalysis.

His scaling methods conceptualized these mental forces as abstract linear continua, objectively measured on numerical scales, and their interrelations expressed as mathematical formulations. Thurstone's sweeping advance, the greatest single achievement rationalizing social experience since the Enlightenment, opened the door to a new science of mind, then stalled when he inexplicably succumbed to factor analysis. The ensuing dark cloud obscured both his measurement and psychology, costing him the momentum to advance psychology to an objective science. In 1954, at the end of his career, he expressed surprise over all the attention received by his difficult factor analytic techniques, while his simple measurement methods never became widely popular (1959, 15).

His most important works, those which promised a sound, objective basis for social research occurred within a short time. By the early 1920s, he had thought out the important conceptual issues for a new science which he discussed philosophically in three articles (1919a, 1923, & 1924). Then in 1925 he explained his new scaling method, quickly followed in 1927, 1928, and 1929 with clarification and elaboration. By 1930, it was over. His shift to primary mental abilities entangled him for years in methods incompatible with absolute scaling. The dark cloud drifted over 20th century social science as factor analysis fatally aroused the naïve enthusiasm of social researchers everywhere.

Volumes could be written about Thurstonian psychometrics: its central features, empirical benefits, and implications for advancing social research. Of course this story would start with his decisive rejection of classi-



Nikolaus Bezruczko

Nikolaus Bezruczko is originally from Linz, Austria and immigrated to the U.S. as an infant. He graduated from MESA, University of Chicago in 1990 and has published widely in child development, aesthetics, psychometrics, and education. He currently lives with wife Ambra Borgognoni Vimercati and stepdaughter Alice in Chicago and Rome, Italy. Favorite recreation is summer hiking in South Tyrolean Alps.

cal psychophysics, as well as raw scores and mental ages. Inconsistencies between Weberian and Fechnerian methods, limen determinations, and JND estimation instability are examples of psychophysical concepts Thurstone considered worthless to social research. To make this methodology meaningful, he needed to reconceptualize psychophysics. Instead of collecting perceptions of lifted weights and constructing a scale with physical units, he would identify distances between mental stimuli based on observer agreement with opinion statements using Fechnerian magnitude estimation methods. Then all he needed was a procedure for transforming ordered proportions into scale values and computing their error distributions. He would project mental structures on linear continua and model their quantitative properties with normal probability functions. Other improvements were also necessary, such as shifting from the method of equal appearing intervals to paired comparison, but the decisive step was to conceptualize a response continuum in terms of social objects such as attitude, opinion, or preference judgments. His ideas, however, were strange to psychologists and social researchers, and Thurstone faced enormous resistance and hostility. He tried to convince skeptics that subjective units were not only sensible and necessary, but easily estimated by selecting an arbitrary item on a continuum and using its error distribution as the scale unit. "The standard deviation of this dispersion for a standard stimulus could be chosen as a subjective unit of measurement." (1952, 307) His responses to objections included elaborate descriptions of his measurement philosophy in publications which fortunately now provide a detailed record of Thurstone's rationale for psychological measurement. Some main ideas are:

* ***Mental integrity.*** A mental integrity independent of overt behavior underlies the human tendency to engage in particular actions. Thurstone's defiant reaction to empty-headed Stimulus-Response psychology, this concept rationalizes an inferential approach to mental functioning.

* ***Discriminal process.*** An automatic perceptual process sorts the ambient flow of external stimuli to identify those that may be useful to the organism. Thurstone asserted they would show an error distribution on the stimulus continuum reproducing the subjective qualitative experience.

* ***Motive forces.*** A structure of motive forces lies dormant in the mental system. Its provocation by items reveals mental affinity toward particular stimuli and defines a psychological continuum. "They acquire conceptual linearity and measurability in the probability with which each of them may be expected to associate with any prescribed stimulus" (1927b, 51). "To the extent their probabilities of association with stimuli are nearly the same, to that extent will they tend to be adjacently spaced on the imaginary psychological continuum." (1927c, 419)

* ***Arbitrary units.*** Measuring in general is based on an arbitrary unit of measure whose practical usefulness is its linearity. Thurstone applied Fechner's JND technique to estimate unit measure on the subjective continuum.

* ***Absolute scaling.*** Social researchers grate at Thurstone's insight that scaling must be independent of the sample measured and unit of measure. (Many of them are still using raw scores/ratings, percentages, and grade equivalents.) "We have called the method absolute, not in the sense of measurement from an absolute origin but in the sense that the scale is independent of the unit selected for the raw scores and of the shape of the distribution of the raw scores" (1927c, 517). If the associational likelihood between any two points on the continuum "should be affected by the opinion of any individual person or group, then it would be impossible to compare the opinion distributions of two groups on the same base" (1928a, 417).

* ***Parameter linearity.*** "The sum of the subjective separations between the stimulus pairs AB/BC must be equal to the experimentally independent determination of the separation AC. If the continuum is unidimensional, then this simple type of check would establish the fact" (1952, 308). Referring to the additivity axiom in physical measurement, Thurstone presages probabilistic conjoint measurement for nonphysical observations.

* ***Item fit.*** Thurstone was explicit, scale items need both rational and empirical support. "The scaling method should be so designed that it will automatically throw out of the scale any opinion statements which do not belong in its natural sequence" (1928a, 417). Thurstone, however, did not support attempts to establish internal consistency coefficients for this purpose. In general, "correlation procedures constitute an acknowledgment of failure to rationalize the problem and to establish the functions that underlie the data" (1929a , 224). Thurstone was adamant, "correlation coefficients are symbols of defeat" (1929a, 240).

He developed a detailed methodology to apply these ideas. For example, Figure 1 taken from a 1926 article presents possibly the first cumulative item response curve ever published in a social research journal, now a standard presentation method. Figure 2 shows parallel item trace lines defining linear structure, the essential empirical evidence for a numerical variable. Another Thurstone contribution to social theory building is the variable map which positions item by person dynamics in a quantitative graphic structure. He considered the map an essential foundation for psychological theory and provided many examples. Figure 3 is his S and R continua, Thurstone's theoretical justification for generalizing psychophysics to nonphysical stimuli (1927a). In the 1920s, any objective, quantitative representation of social phenomena was an extraordinary achievement. Contemporaries such as Binet, Burt, and Thorndike were pioneering ability and

Figure 1

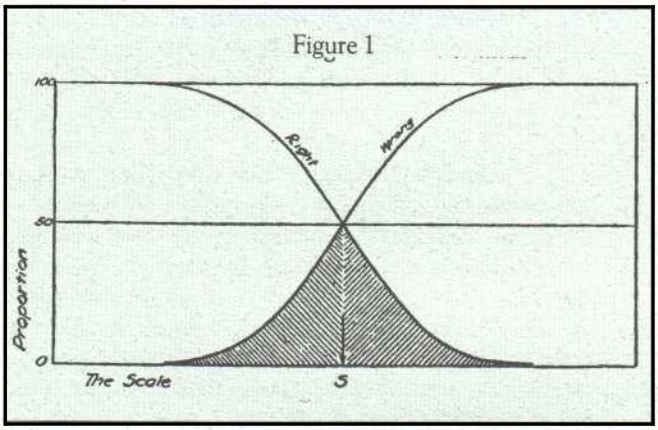


Figure 2

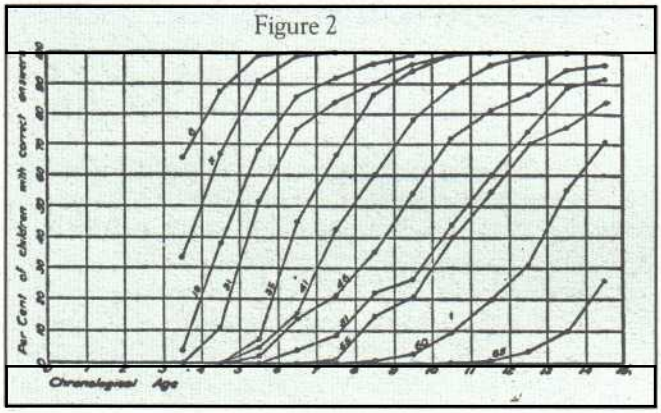
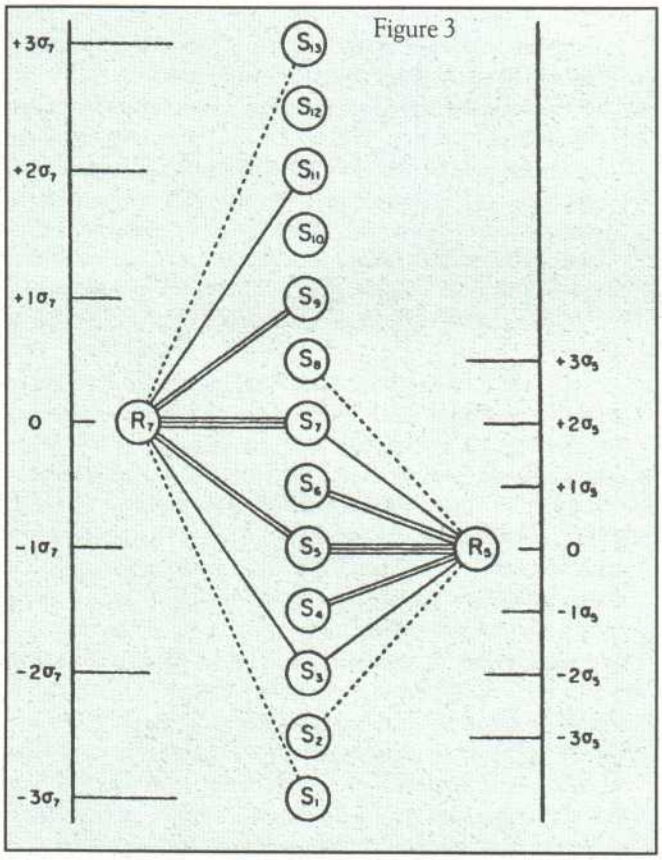


Figure 3



achievement tests, but no one commanded Thurstone's breathtaking view on a new science. Over the next 70 years, his ideas and methods would take on a life of their own ultimately to verify Thurstone's heretical assertion, "Attitudes Can Be Measured" (1928b).

In contemporary social research where hyper-quantification and over-parameterization are endemic, Thurstone is easily dismissed as a historical relic. After all, his whole scaling methodology is based on only two parameters, mean and standard deviation, the scale value and its error distribution. As we all know, the mathematical complications of contemporary social research far surpass Thurstonian methods. The surprise, however, is none of these complicated methods meet scientific rigor. Each of these highly touted methods (multidimensional scaling, cluster analysis, and so on), on close examination, suffers from critical defects that destroy its objectivity, generality, and simplicity. All of them obscure the person in data aggregation. While results are sometimes interesting, they are essentially descriptive techniques about specific samples. None offer any scientific advantages over Thurstonian measurement.

Newton's expression, "If I have seen farther than others, it has been by standing on the shoulders of giants" is appropriate here. The evolution of scientific methodology through Fechner to Thurstone and their successors, carries on an intellectual tradition over 4,500 years old as seen by the balance scales in Egyptian paintings during the Old Kingdom (Rice, 1990). We can only speculate how earlier cultures handled measuring issues. We know humans have an innate tendency to compare objects and abstract their differences. When commensurable with numbers and implemented to describe patterns of uniformity in nature, these units enable the scientific thinking responsible for Western civilization. Separating perceptual units from the observer and re-expressing their quantitative properties numerically is the milestone in human history underlying all abstract sciences. Commerce and its evolution into economics, for example, established social science. The failure of contemporary social research to continue this scientific methodology is responsible for its dismal record in the 20th century. Instead of modeling universal patterns, social research remains limited to fragmented, and inconsistent patterns of testimony, hardly scientific, generally failing to meet even minimal standards of replication or generality. (Some evidence suggests social research has degenerated to cult status, that is, dominated by obtuse methods which are only accessible to high priests yet without any clear relation to constructing scientific knowledge about human behavior.) The current absorption of social research by the physical and biological sciences is a commentary on this failure.

Thurstone provided the architecture for a new science of mind, as well as the foundations for a nonphysical measuring system: an objective framework in which to conduct scientific thinking. His key ideas are continuity, order,

and variability. Continuity is the continuum underlying observations, order is the comparisons among items, and variability is the metric of precision. Georg Rasch, in turn, advanced objectivity by separating the ability and difficulty parameters. This achievement liberates social units from the confinement to standard deviates of arbitrary population means, and constructs a pure mathematical abstraction, a measured difference between ability and difficulty on an infinite continuum. Ben Wright advanced the framework even further by developing tests of statistical fit to detect departures of experience from the abstraction and improve precision and validity. Because this information about persons and items clarifies the dimensionality underlying a scale, it succeeds in eliminating the original motivation to develop factor analysis. Together they establish a measurement trilogy for the 20th century.

Biographical information concerning Louis Thurstone is documented in several sources (Guilford, 1957; Wood, 1962; Thurstone, 1952; see also Gulliksen, 1968). Thurstone was born in Chicago in 1887 to native Swedes, the Thunström family, who changed their name to Thurstone to accommodate American prejudice against foreigners. As a child, he was interested in music reinforced by his musician mother. As a teenager, he became interested in trigonometry and in college published an equation for trisecting any angle (1912). In 1912, he graduated from Cornell University with a mechanical engineering degree and immediately went to work for Thomas Edison in Orange, New Jersey (recruited after demonstrating his model of a nonflickering movie projector). In 1914, he started graduate school in psychology at the University of Chicago. While completing a learning function thesis, he went to Carnegie Institute of Technology in the Department of Applied Psychology. Thurstone returned to the Chicago Department of Psychology in 1924 where he founded the Psychometric Society and the journal *Psychometrika*. (Thurstone spoke on factor analysis to the Sigma Xi Society in spring 1948. After his talk, Ben Wright, then studying physics went to see Thurstone and learned from him his shortcut method for doing factor analysis by hand.) In 1952, Thurstone retired from the University of Chicago and moved his psychometric laboratory to the University of North Carolina. (References available on request.)

REFERENCES

Guilford, J. J. (1957) Louis Leon Thurstone: 1887- 1955. Biographical Memoirs, Volume 30. New York: Columbia University Press for the National Academy of Sciences.

Gulliksen, H. (1968). Louis Leon Thurstone: Experimental and mathematical psychologist. American Psychologist, 23, 716-80.

Rice, M. (1990). Egypt's Making: The Origins of Ancient Egypt 5000-2000 BC. London: Routledge.

Wood, D. A. (1962). Creative Thinker, Dedicated Teacher, Eminent Psychologist. Princeton, New Jersey: Educational Testing Service.

Thurstone, L. L. (1912). A curve which trisects any angle. Scientific American, 73, 259-261.

Thurstone, L. L. (1919a). The anticipatory aspect of consciousness. Journal of Philosophy, Psychology, and Scientific Methods, 16, 561-568.

Thurstone, L. L. (1919b). A scoring method for mental tests. Psychological Bulletin, 16, 235-240.

Thurstone, L. L. (1923). The stimulus-response fallacy in psychology. Psychological Review, 30, 354-369.

Thurstone, L. L. (1924). Contributions of Freudism to psychology: Influence of Freudism on theoretical psychology. Psychological Review, 31, 175-183.

Thurstone, L. L. (1925). A method of scaling psychological and educational tests. Journal of Educational Psychology, 16, 433-451.

Thurstone, L. L. (1926). The scoring of individual performance. Journal of Educational Psychology, 17, 446-457.

Thurstone, L. L. (1927a). Psychophysical analysis. American Journal of Psychology, 38, 368-389.

Thurstone, L. L. (1927b). A mental unit of measurement. Psychological Review, 34, 45-423.

Thurstone, L. L. (1927c). The unit of measurement in educational scales. Journal of Educational Psychology, 18, 505-524.

Thurstone, L. L. (1928a). The measurement of opinion. Journal of Abnormal and Social Psychology, 22, 415-430.

Thurstone, L. L. (1928b). Attitudes can be measured. American Journal of Sociology, 33, 529-454

Thurstone, L. L. (1929a). Theory of attitude measurement. Psychological Review, 36, 222-241.

Thurstone, L. L. (1929b). The mental growth curve for the Binet Tests. Journal of Educational Psychology, 20, 569-583.

Thurstone, L. L. (1952). (Boring, E. G., Langfeld, H.S., Werner, H., & Yerkes, R,M., Eds.) A History of Psychology in Autobiography. 295-321., Worcester, Massachusetts: Clark University Press.

Thurstone, L. L. (1954). The measurement of values. Psychological Review, 61, 47-58.

Thurstone, L. L. (1959). The Measurement of Values. Chicago: University of Chicago Press.

Thurstone, L. L. & Chave, E. J. (1929). Theory of Attitude Measurement. Chicago: University of Chicago Press.

(See Mark Stone's article "Thurstone's Crime Scale Re-Visited" on page 53. Editor)