

昔の評定尺度 —19 世紀初めから 20 世紀初めまで—

椎名乾平 (早稲田大学)

キーワード: 評定尺度, 歴史, 19 世紀

椎名 (教心, 2019) は, Likert 尺度 (1932) 出現以前の 20 世紀初頭において, 既に多様な評定尺度が使用されていたのを示した。本稿はこれより前の時代 (19 世紀初めから 20 世紀初頭) に使用されていた評定尺度についての調査結果を紹介する。以下の諸評定尺度では, Likert 尺度のような目盛直線を用いず, 数値か言語ラベルでの回答を求めている。

英国生物統計学派 (Galton と Karl Pearson)

心理学的な研究での最初の評定尺度使用者を Galton (1880, 1883) とする論者は多い。この尺度は心像の明瞭性を尋ねるものである。Pearson (1906) の評定尺度はよく言及されるが, Pearson (1903) の方が早く, さらにそれらの原型である Pearson (1898) には, 頭の良さについての 6 段階尺度, すなわち, 1. Very dull, 2. Slow dull, 3. Slow, 4. Slow intelligent, 5. Intelligent, 6. Quick Intelligent を自分の学生の評価に用いているという記述がある。

実験美学

Fechner が祖とされる実験美学でも, 評定法が使用されている。Fechner (1871) の選択法 (Methode der Wahl) では, 評価対象の刺激群を同時提示し, その中から心地良い刺激, 心地良くない刺激を選択 (複数選択可) させるという手続きが取られている。三段階の評定と言えないことはないが, 所謂評定法とは発想の異なる手続きである。しかし, その後の Major (1895, 色彩嗜好の研究) の the method of single stimuli では 7 段階評定, すなわち 1: very pleasant, 2: moderately pleasant, 3: just pleasant, 4: without affective tone, 5: just unpleasant, 6: moderately unpleasant, 7: very unpleasant が用いられている。Martin (1905) の漫画の面白さの研究では 5 段階評定が使用されており, Keith (1906) は同様の尺度を "numerical scale" と言い切っている。さらに Johnston (1906, p.160) には, "The scale of numbers from 1 to 7 served in the traditional way to indicate approximately the hedonic value of the feeling-tones, 1 signifying highest degree of pleasure, 2 very pleasant, 3 slightly pleasant, 4 indifferent, 5 slightly unpleasant, 6 very unpleasant, and 7 the highest degree of unpleasantness." とある。この尺度が 1906 年時点で既に traditional と形容されているのは注目値する。

精神物理学

精神物理学的測定と評定尺度は同じとは言えないが, 中には評定尺度と呼べるようなものもある。Hipparchus (BC190-120) が作成した星の明るさの等級 (magnitude) は 6 段階尺度であり, 原文は消失したがブトレマイオスのアルマゲストに載っている。このシステムは 19 世紀には Herschel 親子 (William Herschel, John Herschel) によって精緻化されており (ブットマン, 2009), Fechner (1860) の Elemente でも検討されている。

Titchener (1909) は, 気象学者の素朴精神物理学とでも言うべき気象 (風, 温度, 湿度) の評定尺度について紹介している。特に, Beaufort 風力階級 (13 段階) は海事において 1805 年に考案され 1838 年に英国海軍に公式採用され民間にも広がり, 現在でも使用されている。いくつかの状況証拠より Galton がこの風力階級を知っていたのは確実と思われる。そこで Guilford (1936, p.264) では "Galton's own early associations with the infant

science of meteorology may have acquainted him with the possibilities of rating methods in tackling the problems of human nature." と推測している。

教育評価

Greenwich Hospital School の校長であった, Reverend George Fisher は 1830 年代に 20 段階の尺度を使用し, これは Scale Book と呼ばれていた (ただし原本は失われている)。このことは Ayers (1918) で言及されている。Fisher は海軍の従軍牧師で天文学も修め北極探検に参加した経歴を持つ。また, 勤務校は海軍と関係する学校である。そこで Fisher も Beaufort 風力階級を知っていたのは確実と思われる。Fisher の尺度は先駆的だが後年にどの程度影響を与えたのかは不明である。

現在も使用されている数字・記号による成績段階評価 (A, B, C, ... 等) も相当早くから用いられていたようである。Meriam (1905) は教師教育の研究であるが, その中で Examinations are marked 98%, 86%, 37%, etc.; or by letters A, B, B-, C, C-, D, etc.; or by numbers 1, 2, 3, etc.; or by words "excellent," "good," "poor," etc. とあり, 当たり前使用されているという (p.58)。心理測定学の影響を直接受けていないこのタイプの「自然発生的教育評価法」は, Fisher の例もあるとおり心理学より古いのであろう。心理学の見地から見直そうという研究もある (例えば Finkelstein 1913 The marking system in theory and practice)。

順位法と評定法

Rank order 法と Order-of-merit 法はほぼ同意語で順序付けを行う方法である。特に Order of merit は Cattell, J. M. を中心とするコロンビア大学学派での用語で「勲等」という語感が強いと思われる。大正時代にはどちらも「品等法」と訳されることがあった。評定法は同順位を許す順序法なのだと解釈すれば同じものとなる (心理学的プロセスは同じとは言えないだろう)。例えば, 順位法と評定法の間型型と呼べる様な方法に Hollingworth (1911, 1913) の Group method がある。この方法では刺激をカードに書き, そのカードを与えられた基準で 10 個の (順序付き) pile に分類する。こうなると Likert 法とほとんど区別がつかない。

まとめ

1800-1910 年ぐらいの評定尺度は目盛直線がないのを除けば, 現在のものとあまり変わらないように思える。一方, 参照すべき強い基準のあるもの (Beaufort), 段階について詳しい情報を付加するもの (現代の rubric のようなもの) が多いように思われる。

主要文献

Buttmann 2009 星を追い, 光を愛して; Fechner 1871 Zur experimentalen aesthetik; Guilford 1936 Psycho-metric methods; Hollingworth 1913 Experimental studies in judgment. Archives of Psychol., 29.; Johnston 1906 The combination of feelings. Harvard Psychol. Studies, 2.; Major 1895 On the affective Tone ... Amer. J. of Psychol.; Meriam 1905 Normal school education and efficiency in teaching. Teachers college, Columbia univ.; Titchener 1909 The psychophysics of climate. Amer. J. of Psychol.; Pearson 1898 On a scale of intelligence in children. J. of Education, 509-10.