Measurement 71 (2015) 46–55

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

# Counting and quantification: Comparing psychometric and metrological perspectives on visual perceptions of number

L.R. Pendrill [a,*], William P. Fisher Jr. [b,c]

[a] SP Technical Research Institute of Sweden, Measurement Technology, Box 857, SE-50115 Borås, Sweden
[b] BEAR Center, Graduate School of Education, University of California, Berkeley, CA, USA
[c] LivingCapitalMetrics Consulting, Sausalito, CA, USA

ARTICLE INFO

*Article history:*
Received 9 June 2014
Received in revised form 26 March 2015
Accepted 8 April 2015
Available online 23 April 2015

*Keywords:*
Human
Response
Health care
Metrology
Instrument
Psychometrics

ABSTRACT

In gaining a better understanding of how to characterise human response, essential to improved person-centred care and other situations where human factors are crucial, recent work has attempted to link metrological (resolution, classification effectiveness) and psychometric (Rasch) characterisation of Man as a Measurement Instrument. The present work offers a more detailed account of these investigations following our first preliminary conference report, continuing a study of elementary tasks, such as counting dots, where one knows independently the expected value because the measurement object (collection of dots) is prepared in advance. The analysis is compared and contrasted with recent approaches to this problem by others, for instance using signal error fidelity and loss functions. Independent sources of measurement uncertainty, such as under-estimation of scores, are distinguished from separate estimates of task challenge and individual counting ability, and accounted for in estimates of reliability of the various measures.

## 1. Introduction

The reliable characterisation of the human measuring instrument [1], be it with the five senses or the full physiological, mental, cognitive and behavioural richness of human perception, is essential in many applications. Some of these include enhancement of various human functions, machine learning [2] to assist in mining the ever increasing amounts of information available in society, or aiding a disabled, ill or elderly person to cope better with everyday tasks [3], to name a few. Quantities of concern are not merely technical but also more human, such as comfort, pleasure, and beauty [4].

Bearing in mind that formulation of metrological concepts commensurate with those established in traditional engineering is as yet in its infancy in perceptual contexts [5–8], we have initiated work attempting to link metrological (resolution, classification effectiveness) and psychometric (Rasch) characterisation of Man as a Measurement Instrument, as briefly reported in a conference proceedings [9]. The present work offers a more detailed account of these investigations, continuing our study of elementary tasks, such as counting dots, where one knows independently the expected value because the measurement object (collection of dots) is prepared in advance. Two key aspects of quality-assured measurement – traceability and uncertainty – must be kept in focus when metrologically characterising Man as a Measurement Instrument.

Some method of metrological traceability to invariant unit standards for measurements based on ordinal observations is needed when the ability of a person to perform a task of classifying an entity of given reference level is to be determined. Patient health, for example, is increasingly rated in health clinics on ordinal scales linearized

---

* Corresponding author. Tel.: +46 767 88 54 44.
*E-mail addresses:* leslie.pendrill@sp.se (L.R. Pendrill), wfisher@berkeley.edu (W.P. Fisher Jr.).

http://dx.doi.org/10.1016/j.measurement.2015.04.010
0263-2241/

via a log-odds transformation, and appropriate treatment is decided by comparing the actual ratings with corresponding patterns of typical health ratings for similar patients from earlier studies. The comparability of such ratings has to be reliable to a sufficient degree of accuracy if the patient is to be treated appropriately. When and where such accuracy can be regularly obtained, the observational framework could be redesigned to omit the observed ordinal scores and to incorporate a metrologically traceable unit [7,10,11].

The second aspect of quality-assured measurement – namely, measurement uncertainty and reliability – also presents some challenges where observations are made with Man as a Measurement Instrument. In particular the usual tools of statistics, such as for calculation of the mean and standard deviation, needed when expressing uncertainty, cannot generally be applied to scores obtained with questionnaires and similar instruments, since a lack of an invariant unit renders uninterpretable the location and dispersion of qualitative measurements on ordinal scales [12]. At the same time, measurement uncertainty, since it reflects the quality of measurement, also provides a measure of the ability of human to perform as an instrument.

## 2. Rationale for a potential expansion of the metrological framework

An inquiry into the viability of a uniform unit of physical functioning in medical rehabilitation compared results produced across four different measurement instruments applied to eleven different samples [7]. The heuristic model employed producing each set of results demands estimates "not affected by the abilities or attitudes of the particular persons measured, or by the difficulties of the particular survey or test items used to measure". Fit to this model allows parallels to be drawn between psychometric concepts of invariance and equating, on the one hand, and metrological concepts of traceability to repeatable and reproducible unit standards, on the other. This need for more rigorously defined and more widely distributed measures is implicated by the trend in health care associated with a shift from "local economies of disease-crisis management to regional, national, and international economies of population-based, preventive health management". As demand for proactive prevention displaces reactive responses, it is virtually inevitable that continuing growth in the speed and networking reach of computational tools will propel invariant measurement into significant new roles supporting accountability and comparability in health care.

Human beings inevitably play critical roles in measurement [13]. From the perspective of engineering, the operation of any measurement system requires calibration, data acquisition, and data presentation [14]. Recent studies in psychophysical scaling [15] combine psychometric and engineering perspectives, relating perceptual intensity to stimulus intensity by explaining the Weber–Fechner law in terms of signal error fidelity [16]. Other studies describe adaptation by biological sensory systems in terms of the costs of perceptual task errors [17].

Another approach to linking engineering and psychometric conceptualizations of measurement systems is suggested by psychological measurement models introduced by Rasch [18–20]. When a human is instrumental to the performance of elementary tasks – such as counting dots [21,22] – person abilities relative to the degree of challenge posed by different tasks can be expressed in terms of measures invariantly located and dispersed on an interval scale [23,24]. The ability to perform the task can be calibrated and measurement uncertainty can be assessed in this context with the special advantage of independent advance knowledge of the measurement object's true value (a given number of dots).

## 3. Grounding measuring in counting

Our first preliminary brief report in a conference proceedings [9] highlighted the suitability for our research of previous studies by others [22] concerning the counting ability of the Mundurucu, an Amazonian indigenous people with little access to Western-style educational resources and where counting above the number five is often a challenge. Research investigating the conceptual link between number and spatially distributed dots had already suggested that the Mundurucu intuitively employ a logarithmic transformation of impressions of varying amounts (Fig. 1), meaning that "larger numbers require a proportional larger difference in order to remain equally discriminable" (Weber's law) [22]. In the present work, we extend the analysis of Dehaene et al. [22], taking advantage, as they, when attempting to characterise human response, of the conceptual simplicity where one knows independently the expected value (the number of dots). Our aim is to explore further the link between metrological (resolution, classification effectiveness) and psychometric (Rasch) characterisation of Man as a Measurement Instrument.

Fechner was among the earliest to note the contrast between the linearity of measures and the nonlinearity of intuited impressions for the human senses, a contrast of ongoing interest in neurological research [25]. Because a fairly constant degree of imprecision is maintained across several orders of magnitude, the "Gaussian tuning curve" serves, in effect, as a kind of internally embodied sensory slide rule [22]. By identifying and describing the logarithmic proportionality of sensations and stimuli, Fechner connected physical experience with linear geometry in a way that set the stage for Thurstone, and, later, Rasch, to refocus human measurement away from its previous preoccupation with purely psychophysical phenomena to broader concerns with psychological, economic, and social phenomena [26,27].

It is important to note the deeper connection here that real things, like the sides of triangles, rocks, or human behaviours, are never identical, and so do not ever conform perfectly with expectations formed on the basis of a mathematical formulation of a scientific law or measurement model. Measurement, whether of counting ability or of mass or temperature, requires abstract invariant units that physically cannot correspond directly with empirical