Chapter 3
Ensuring Traceability

This chapter deals with traceability and comparability: the first of the two major hallmarks of metrology (quality-assured measurement). The second hallmark—uncertainty—is covered in the final chapters of the book.

In Fig. 3.1 is recounted a parable illustrating the concept of measurement comparability, in this case of the physical quantity of time, but the meaning should be universal.

A retired sea-captain (a) on an island takes his time from the watchmaker (b) in town only to find out that the watchmaker uses the sea-captain's cannon shots at 12 noon each day to set his own clocks! (Attributed to Harrison (MIT) by Petley 1985). Examples of this kind of circular traceability in measurement are more common than one would hope. The parable captures the essence of the concept of trueness, that is, what defines being 'on target' when making repeated measurements in the 'bull's eye' illustration of Fig. 2.2?

Reliable measurement results are important in almost every aspect of daily life. Comparability of entity properties is critical to global trade between producers and consumers; provides for the interoperability and exchangeability of parts and systems in complex industrial products; is essential in the synchronisation of signals in communication systems and provides a fair and quality-assured basis for measurement in the environmental, pharmaceutical and other chemical sectors. Metrological traceability through calibration enables the measurement comparability needed, in one form or another, to ensure entity comparability in any of the many fields mentioned in the first lines of Chap. 1.

Despite its importance, international consensus about traceability of measurement results—both conceptually and in implementation—has yet to be achieved in every field. Ever-increasing demands for comparability of measurement results needed for sustainable development in the widest sense require a common understanding of the basic concepts of traceability of measurement results at the global level, in both traditional as well as new areas of technology and societal concern.

The present chapter attempts to reach such a consensus by considering in depth the concept of traceability, in terms of calibration, measurement units and standards

L. Pendrill, *Quality Assured Measurement*, Springer Series in Measurement Science and Technology, https://doi.org/10.1007/978-3-030-28695-8_3

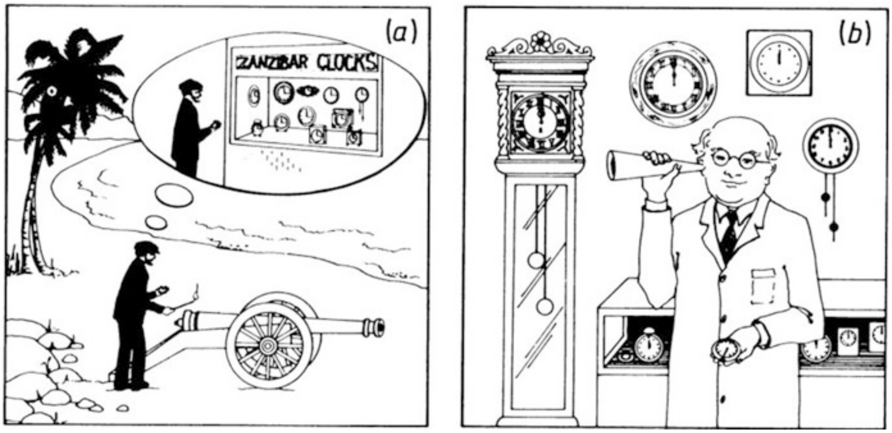

Fig. 3.1 Zanzibar effect (Reproduced with permission Petley 1985)

(etalons), symmetry, conservation laws and entropy, in a presentation founded on quantity calculus. While historically physics has been the main arena in which these concepts have been developed, it is now timely to take a broader view encompassing even the social sciences, guided by philosophical considerations and even politics. At the same time as the International System of Units is under revision, with more emphasis on the fundamental constants of physics in the various unit definitions, there is some fundamental re-appraisal needed to extend traceability to cover even the less quantitative properties typical of measurement in the social sciences and elsewhere.

3.1 Quantity Calculus and Calibration

What use can be had of quantity calculus when seeking measurement comparability needed, in one form or another, to ensure entity comparability in any of the many fields of application?

With much of measurement developed over many years in the context of physics and engineering, there has been a long-standing connexion with applied mathematics: 'The object of measurement is to enable the powerful weapon of mathematical analysis to be applied to the subject matter of science' according to Campbell (1920). The renowned theoretical physicist Feynman (2013) went as far as to state: 'Experiment is the *sole judge* of scientific "truth."'

In developing theory in quantum mechanics almost one hundred years ago, Dirac [§5, p.15 1992) wrote: 'In an application of the theory one would be given certain physical information, which one would proceed to express by equations between the mathematical quantities. One would then deduce new equations with the help of the axioms and rules of manipulation and would conclude by interpreting these new equations as physical conditions. The justification for the whole scheme depends, apart from internal consistency, on the agreement of the final results with experiment'.

Quantity calculus, as developed over much of the twentieth century, is described echoing Dirac from the metrological point of view as: 'mathematical relations between quantities in a given system of quantities, independent of measurement units' (JCGM 200: 2012, VIM §1.22). Quantity calculus, according to the review of de Boer (1994/5), has its roots in the mathematical formalism of theoretical physics dating from the pioneering work of Wallot (1926). At the same time, it cannot be said that there exists yet a general consensus about an axiomatic foundation for quantity calculus (de Boer 1994/5; Raposo 2016). This incompleteness is becoming of increasing concern as interest in quality-assured measurement spreads to new sectors such as the social sciences.

3.1.1 Quantity Concepts

In defining quantity calculus, an hierarchy of concepts is traditionally established in the order: Kind of quantity; Quantity; Value of quantity. Fleischmann (1960) gave some examples of the concepts at the different hierarchy levels (my translation from the German):

- 'For each single feature, every kind of quantity can have an infinite number of quantities (characteristic quantities) of the feature.
- When the quantity is not specific, one refers to a 'Quantity' or 'General quantity'. For example: the quantity electric voltage V can take the value 1 V or the value 20 V, etc.
- An 'entity quantity', in the terminology of Fleischmann (1960),[1] is the case where a general quantity can be associated with an entity (object). For example, the effective voltage and maximum voltage (for sinusoidal alternating current) are distinct entity quantities.
- Quantity values are specific values of quantities. For instance in the area of electrical voltage, one can have quantity values 5 mV, 30 V, etc. Examples of quantity values which belong to different kinds of quantity are 10 s and 3 cm. Quantity values are associated with a definite quantity.'

The superordinate concept of 'kind of quantity' in quantity calculus (as in Fleischmann 1960) is used to collect together quantities by kind or character. Quantities themselves and relations between them, that is, in the 'entity (or product) space' (Chap. 1), irrespective of whether they are measured or not, come next in the conceptual hierarchy. Thereafter, aspects in the measurement space

[1]German: '*Sachgrösse* = Objektgrösse (mit Objektbindung behaftete Grösse). Sie hat Quantität, die aber unbestimmt bleibt, sie hat Sachbezug (Objektbezug)' Fleischmann (1960).

(Chap. 2), such as quantity values, can be introduced subsequently, as required (Sect. 3.2).

Which of these various groupings are chosen depends on what is needed and meaningful in each field of application when in general communicating measurement information.

In full generality, relations in quantity calculus can be formed between different kinds of quantity and quantities in a more abstract sense. Consideration of 'general' quantities is much the domain of the physicist (Pendrill 2005) where relations (laws of Nature) among such quantities (which also give the corresponding relations among the measurement units associated with them—see Sect. 3.2) are fundamentally and universally applicable, *irrespective* of particular objects (as for instance in Newtonian mechanics as applied to all bodies, from microscopic and cosmological scales). Measurement is a necessary action in physics, but the main interest is in understanding the universe.

Physical quantities possess some remarkable properties which enable correspondingly remarkable possibilities for metrologically traceable measurements: not only can the results of measurements of a particular quantity be compared, but also measurements of *different quantities* can show a degree of comparability (Sect. 3.6.1). Relations between physical quantities can be expressed mathematically with equations that express laws of Nature or define new quantities.

Quantity calculus can be used to highlight the interesting distinction between:

- a physical law, e.g. force $F = m \cdot a$ (Newton's second law, if the mass, m, is constant) relating different quantities is universal; applicable at all scales from the microscopic to the cosmological,
- those indirect measurements where only an empirical 'recipe' is used, e.g. an engineering expression relating a number of different properties for a particular object, which has local validity but only limited universality. An example is an expression for the volume, V, of combustion chamber above a combustion engine piston $V = \varepsilon\ V_k$, where V_k – volume of combustion chamber when piston is in upmost position; ε – compression ratio (Kogan 2014).

The remarkable properties in physics are of course not necessarily shared by quantities in other disciplines (Pendrill 2005; Nelson 2015), which is a key issue of course in the context of the present book, aiming to give a unified presentation of quality-assured measurement across the social and physical sciences. In the last column of Table 3.1, an example—for the unit of length—is given to illustrate the various ways of expressing the unit at each level of the quantity calculus hierarchy.

Table 3.1 Comparing concepts in information theory and quantity calculus

Quantity calculus	Information theory	Examples: SI unit of length, metre, symbol m
Nature of quantity	*Effectiveness*—'changing conduct': relationship between signs of communication and actively 'improving' the entities they stand for (Weinberger 2003)	Quality of cloth products sold by metre length
Kind of quantity	*Pragmatic*—'utility': relationship between signs of communication and their utility (value, impact)	Length of cloth costing 10€/m
Entity quantity	Object quantity (object-bound quantity). It has quantity, but remains indefinite; it has reference to reality (object reference) (Fleischmann 1960)	Length of cloth
Quantity	*Semantic*—'meaning': relationship between signs of communication and entities they stand for	Distance travelled by light in 1/c seconds ('explicit unit' definition, CGPM 2019)
Value of quantity	*Syntax*—'signs': relationship among signs of communication such as numbers	Defined by taking fixed numerical value of speed of light in vacuum c to be 299 792 458 when expressed in unit m/s, where the second is defined in terms of $\Delta\nu_{Cs}$ ('explicit constant' definition, CGPM 2019)

3.1.2 Introducing Measurement and Calibration. Separating Object and Instrument. Restitution

At the lowest level of the hierarchy of concepts of quantity calculus, mathematical relations can be formulated specifically between quantity values, firmly in the realm of measurement.

In general, the value of an item attribute δ (e.g. a level of challenge of a particular task or the mass of a weight) differs from the 'true' δ', by an error ε_δ:

$$\delta = \delta' + \varepsilon_\delta \tag{3.1}$$

For comparability, it is necessary to communicate information about the extent to which the item attribute value is in error. As explained in Sect. 2.1, it is necessary (and challenging) to distinguish between actual product value and an apparent value distorted by limited measurement reliability.

The calibration process—that is, determining the error (with respect to the 'true value' depicted as the bull's eye at the centre of the target shown in Fig. 2.2)—often identifies the attribute of a particular object (called a measurement standard or etalon) as a known reference with which other attributes (and their errors) can be referenced.

The international metrology vocabulary defines (JCGM VIM 200:2012):

5.1 *measurement standard* etalon

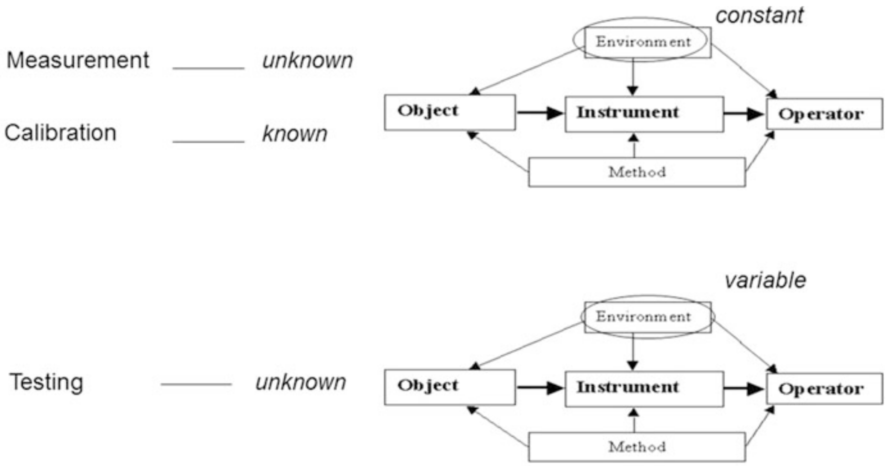

Fig. 3.2 Measurement, calibration and testing under different conditions of a measurement system

> realization of the definition of a given quantity, with stated quantity value and associated measurement uncertainty, used as a reference
>
> NOTE 1 A 'realization of the definition of a given quantity' can be provided by a measuring system, a material measure, or a reference material.
>
> NOTE 2 A measurement standard is frequently used as a reference in establishing measured quantity values and associated measurement uncertainties for other quantities of the same kind, thereby establishing metrological traceability through calibration of other measurement standards, measuring instruments, or measuring systems.

Calibration of a measurement system for instance consists of applying a stimulus of known value (from the standard) to the input of the measurement system, and then determining the extent to which measurement error can be associated with every element of the measurement system (Fig. 2.5).

The process of calibration is illustrated in Fig. 3.2 in comparison with the related but distinct operations of testing a measurement system as well as making measurements with a (calibrated) measurement system to determine an unknown quantity.

Calibration is defined in the international vocabulary (VIM §2.39) as:

> Operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication.

In terms of measurement system analysis, it is important to distinguish between the quantity value which is the measured response of the instrument to a known stimulus, and the value of the entity as estimated with the process of restitution. We interpret the 'measurement result from an indication' in the VIM calibration definition as referring to the latter.

A calibrated measurement system is of course a pre-requisite in a usual measurement situation where one uses the system to determine the unknown stimulus from the measurement object. The restitution process which converts the system response into the stimulus value requires a known sensitivity of the measurement instrument (Sect. 2.4.5). In the case where the item (entity) is acting as a calibration standard (etalon), then the error ε_δ (Eq. (3.1)) is known (from a previous calibration process). Observing the response of the measurement system being calibrated and tested with the known stimulus allows determination of in principle all characteristics associated with the measurement system (listed in Table 2.4). Among the most frequently determined characteristics when calibrating a measurement system are the sensitivity, K, of the instrument and the bias, b. For instance, $b_{\text{cal}} = b - \varepsilon_\delta$. If both K and b are known from calibration and can be assumed to be stable enough[2] that they remain substantially unchanged in value on later usage of the calibrated measurement system, then the process of restitution yields an estimate of an unknown stimulus in the simplest case with the expression

$$S = \frac{R - b_{\text{cal}}}{K_{\text{cal}}} \tag{2.9}$$

If the ultimate aim of a measurement is not only to determine an error, but to make a decision of conformity for an entity, then restitution will yield estimates, δ, of the respective properties of interest—for example, measurements in the social sciences such as of the difficulty of a task; the quality of a service or beauty of the portrait. As described earlier (Sect. 2.4.5), restitution for categorical responses (e.g. of a human being's response as an 'instrument' to a stimulus) is made in terms of P_{success}, that is, the probability of making a correct categorisation and the Rasch measurement model, mentioned in Sect. 1.2.3 and Eq. (1.1) can be applied:

$$S = z = \theta - \delta = \log\left[\frac{P_{\text{success}}}{1 - P_{\text{success}}}\right] \tag{1.1}$$

In contrast to physical measurement systems, the measurand (quantity to be assessed in the social sciences, such as the degree of quality of care, δ)—which is the stimulus S of the measurement object characteristics—is estimated through restitution of the measurement system response R of instrument (sensitivity K) in terms of a performance metric (how well the task is performed or how well the quality of a service is rated) using Rasch invariant measurement theory, rather than an inversion merely based on measurement error.

Measurement is a 'concatenation of observation and restitution' (as recalled by Bentley 2004; Sommer and Siebert 2006; Rossi 2014).

A full picture of the measurement process when Man acts as a measurement instrument can be given, as in Fig. 3.3, presenting the process, step by step, from the observed indication (a performance metric, e.g. probability of success, P_{success}, of

[2] Chapter 5 contains a description of how to test these assumptions.

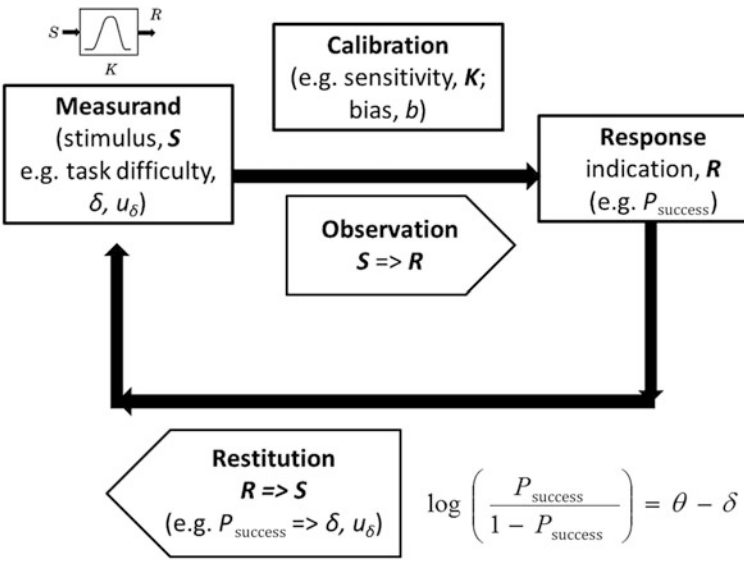

Fig. 3.3 Observation and restitution for performance metrics (adapted from Pendrill 2018)

achieving a task), and restitution with Rasch Measurement Theory, through to the measurand (e.g. task difficulty) in a form suitable for metrological quality assurance.

Irrespective of whether measurements are made in the physical or social sciences, a fundamental requirement for establishing metrological standards (etalons) for calibration (be it of other etalons or measurement systems) is that a separation can be made in the response of the measurement system between the stimulus value—that of the measurand associated with the measurement object—and the characteristics of the measurement system—particularly the sensitivity and eventual bias of the measurement instrument. Without that separation, there is no way of reliably performing the restitution process nor establishing the all-important metrological aspects of traceability and comparability.

In connexion with the introduction of a revised SI (Sect. 3.6.1), the exact terminology to be used to describe quantities, quantity values, and units has been debated in a number of international committees as well as in the research literature (Kogan 2014; Mari et al. 2018). In a first case study (Sect. 3.6.1) at the end of this chapter, we consider how quantum calculus can be invoked to make clear the conceptual difference between quantity and quantity value. In a second case study (Sect. 3.6.2), establishment of metrological standards for calibration in the social sciences (and other 'qualitative' situations) will be considered.

3.2 Units and Symmetry, Conservation Laws and Minimum Entropy

Communicating information about the extent with which the item attribute is in error can be seen as a description of metrological comparability in a purely mathematical sense as in Eq. (3.1). But measurement is not only mathematics and in the present section we consider the additional aspects which guide us when choosing among the various groupings of quantity calculus (Sect. 3.1.1). In particular, what kind of measurement information is being communicated and what is the 'meaning' behind an error ε_δ?

Which grouping of quantity calculus concepts one chooses depends on what is needed in each field of application when communicating measurement information in general. Roberts (1985), in his introduction to measurement with applications in the social sciences, illustrates this with a number of statements (or messages) which appear either meaningful or meaningless (Table 3.2).

Two issues about measurement are raised by Roberts' (1985) examples:

- Measurement has much to do with providing clear, meaningful messages with which to communicate information about what (entity, quality characteristic...) is being measured.
- Measurement units are a key concept in enabling efficient communication.

In this section, these and related concepts such as entropy, symmetry, conservation will be presented which in different ways relate to the efficient and meaningful communication of measurement information. The introduction of meaningfulness of empirical scalings and analyses based on them, while elementary, has not been without statistical controversy where some statisticians perceive the approach as too prescriptive (Velleman and Wilkinson 1993). Further discussion of quantitative and qualitative scales of measurement is given in Sect. 3.5.

Table 3.2 Meaningful messages about measurement

Number of cans of corn in local supermarket at closing time yesterday was at least 10	Meaningful
One can of corn weighs at least 10	Meaningless
One can of corn weighs twice as much as a second can	Meaningful
Temperature of one can of corn at closing time yesterday was twice as much as that of a second can	Meaningless

Adapted from Roberts 1985

3.2.1 Meaningful Messages and Communicating Measurement Information

Regarding measurement as a particular kind of information was briefly eluded to earlier in this book (Sect. 2.2.3) when introducing measurement system analysis (MSA) in terms of a faithful description of the observation process, that is, how measurement information is transmitted from the measurement object, via an instrument, to an observer. Apart from this local communication, it is of course in many cases also required to communicate measurement information as globally—across distances and among different persons—as needed, according to what is meaningful in each field of application (listed at the start of Chap. 1). In principle, the manner of formulating measurement information will depend on which level of the quantity calculus hierarchy one is at (Table 3.1).

When considering a suitable choice of types of quantities in quantity calculus in view of what kind of information is considered important to communicate, Emerson (2008) for instance argues that 'kind of quantity' (Sect. 3.1.1) has to be complemented by adding the distinct concept of 'nature of quantity', and gives examples of two quantities of the same kind, but which are 'meaningless' to compare:

- 'the length of the distance between two rail termini has a different nature to the length of the track gauge of the same railway,
- the height of tide at London Bridge which is the same kind of quantity, but of a different nature, to the height of tide at Washington Bridge'.

Other examples are given in Table 3.1. Emerson (2008) takes such examples of the comparability of kinds of quantities as illustrations of extensive quantities which can be added but only if their 'datum values are not associated with different and immovable places or times'.

Similar discussions may be found in infology (Langefors 1993), where 'information' is tied to the relevance of the knowledge to the decision to be made. Knowledge that has no bearing on the problem at hand will not reduce the uncertainty associated with the problem, and will not be recognised as information. A minimum information element would contain at least three parts: one part referring to the entity informed about; another part which refers to the property of the entity and a third, a 'locator', such as a time reference part in the sentence. For example: 'The temperature of container C was 65 degrees on April 15, 1992 in room X'. In modern information terminology, one could refer to these as conditions for maintaining the integrity of a message on transmission through a communication system—so-called semantic interoperability (Marco-Ruiz et al. 2016 give examples from health informatics). At the pinnacle of the quantity calculus hierarchy Weinberger (2003) identifies the concept of information *effectiveness*, meaning 'changing conduct' in terms of relations between signs of communication having to do with actively 'improving' the entities they stand for.

Transmission of measurement information is of course a specific case of communication with an information system (Pendrill 2011). Be it with a local measurement system (as depicted in Fig. 2.4) or more globally across distances and among different persons, the 'transmitter' (event 'B'), 'information channel' (signal 'B' => 'A') and 'receiver' (message 'A') of a classical information system in the case of a measurement system correspond, respectively, to the measurement object, measurement instrument and observer. The amount of information transmitted from the measurement object to the observer can range from a simple signal through to increasingly 'meaningful' messages, as is captured in four levels of increasing richness in information theory (Weaver and Shannon 1963; Klir and Folger 1988) as given in the first column of Table 3.1. Depending on what kind of meaning is to be communicated, the kind of (measurement) information will fall into one or other of the extended quantity calculus hierarchy (second column of Table 3.1).

3.2.2 Units, Words and Invariance

A useful point of departure for introducing units in measurement is to recognise the analogous role of words in making efficient communication with language. Metrological traceability enables the measurement comparability needed, and calibration (Sect. 3.1.2) involves tracing the measurement standards which, so to speak, embody defined and 'recognisable' or 'meaningful' amounts of the measurement unit.

A classic example illustrating the different information content of the following three messages consisting of an equal number of digits (or bits):

'100110001100'
'agurjerhjjkl'
'this message'

It is obvious to everyone who understands English that the third message conveys more information than the two other messages. In the next section a connexion will be made between the amount of information conveyed by a message and informational entropy, where the latter is a measure of the degree of 'order' in broad meaning in the message contents.

The international metrology vocabulary [VIM §1.9] defines *measurement unit* as a:

> real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number.

Because of the key role played by measurement units, there have to be both clear definitions of each unit as well descriptions of how each unit is 'realised'. To be of any practical use, units do not only have to be defined, but they also have to be realized physically for dissemination. A variety of experiments may be used to realise the definitions—called 'mise en pratique'. Current definitions of the measurement units of the International System (SI) can be found in the SI brochure (CGPM 2019).

The original text by Maxwell referring to measurement units with the expression Eq. (3.2):

$$Q = \{Q\} \cdot [Q] \tag{3.2}$$

is repeated in the respected paper by de Boer (1994/5): 'Every expression of a quantity Q consists of two factors or components. One of these is the name of a certain known *quantity* $[Q]$ of the same kind as the quantity to be expressed, which is taken as a standard of reference. The other component is the number of times $\{Q\}$ the standard is taken in order to make up the required quantity'.[3] We go one step further when pointing out that measurement units—as recognisable 'packets of measurement information'—play a role in communicating meaningful measurement information analogous to words in a language.

An interpretation of Eq. (3.2) is that any measurement of the quantity Q consists—as expressed by Maxwell 'making up'—of displacing the unit and counting how many times it fits in the measured displacement, where 'displacement' is not specifically in length, but in the dimension of interest. An assumption implicit in this procedure is that space is invariant in the dimension of the measured quantity, so that the unit as embodied in a measurement standard does not change on translation. The connexion with measurement units is made by the observation that $Q = \{Q\} \cdot [Q]$ relies on the invariance of the unit quantity upon measurement transformation. Conditions for invariance will be discussed in Sect. 3.2.3.

This interpretation of measurement units in terms of invariance is at a more philosophical and fundamental level than mere engineering. As mentioned above (Sect. 3.1.1), an engineer might see superficial similarities between Maxwell's Eq. (3.2) and for instance and expression for the volume, V, of combustion chamber above a combustion engine piston, $V = \varepsilon\ V_k$, where V_k – volume of combustion chamber when piston is in upmost position; ε – compression ratio (Kogan 2014). However, not all ratios of quantities reflect fundamental symmetries: Presumably in most applications there are more prosaic (albeit essential) factors, such thermal expansion, mechanical properties of piston material, etc., which affect the piston volume and need to be corrected for, and which are usually much larger than the effects of any fundamental symmetry breaking.

A case study of measurement units in the new SI (Sect. 3.6.1) will describe a connexion between measurement units and quantum mechanics, including unitary transformations. A second case study concluding this chapter (Sect. 3.6.2) will deal with measurement units in psychometry, with broader application in measurements in the social sciences.

[3]Note however that there is no explicit reference in Eq. (3.2) to which object/entity is being measured, but rather to a certain kind of quantity.

3.2.3 Symmetry, Conserved Quantities and Minimum Entropy. Maximum Entropy and Uncertainty

How much information that is carried by a word (or a unit, (Sect. 3.2.2)), or any other message, can be measured in terms of the concept of entropy in information theory (Shannon and Weaver 1963). Two distinct contexts where one seeks stationary (minimum or maximum) values of the entropy—that is, the amount of information—can be identified: (1) the best units for traceability are those with most order, i.e. least entropy, as an example of the principle of least action; (2) the change in entropy on transmission of measurement information cannot decrease, thus allowing realistic estimates of measurement uncertainty, in line with the second law of thermodynamics.

Firstly, symmetry—invariance under displacement—is a pre-requisite when forming easily recognised and 'meaningful' words and measurement units. The more symmetric—or ordered—a word is, the lower the entropy. Symmetry, having smaller entropy (most order), most meaning and unchanged on displacement, allows for a more efficient 'packaging' or 'chunking' of the information in a message, so the content is maintained on transmission and is more readily understood by a receiver. Attneave (1954), Barlow (2001), Miller (1956), Schneider et al. (1986) and others addressed how recognisable ('meaningful') patterns can improve communication by exploiting redundancy or 'chunking'. A basic concept in metrology is to choose a measurement unit which has the most order, i.e. least entropy, $H(P)$, associated with the message P. In terms of meaningfulness (Sect. 3.2.1), most information is contained in a message in which the entropy is minimised—which could be interpreted either in terms of the 'simplest' or most 'likely' signal. Meaningfulness is of course determined in each case by what the ultimate aim of the measurement is.

Figure 3.4 exemplifies a number of units of measure which are associated with recognisable patterns reflecting transformational symmetry (low entropy): a unit of time (**t**) for instance found in diverse physical systems (clock, atomic transition, planet, pulsar. . ., as in SI definition of second), where the canonical variable energy (**E**) is conserved under a 'displacement' through a 'distance'.

It is well known that the transformation symmetry sought when defining measurement units is related to conservation of the measured quantity. In seeking suitable systems with which to define and realise measurement units, one can observe that a number of physical quantities are known by experiment to be conserved in an isolated system: the total *energy, momentum, angular momentum* remain constant, whatever and however complex interactions occur within the system. These constancies are consequences of the invariance of mechanical systems under changes of corresponding canonical quantities—*time*, under *length* translation and under *rotation* in space, respectively—together with the principal of least action (Landau and Lifshitz (1976)). The role of entropy and symmetry in measurement is useful, not only in physics but also in the social sciences when seeking metrological standards for measurement. Examples include defining the difficulty of a cognitive

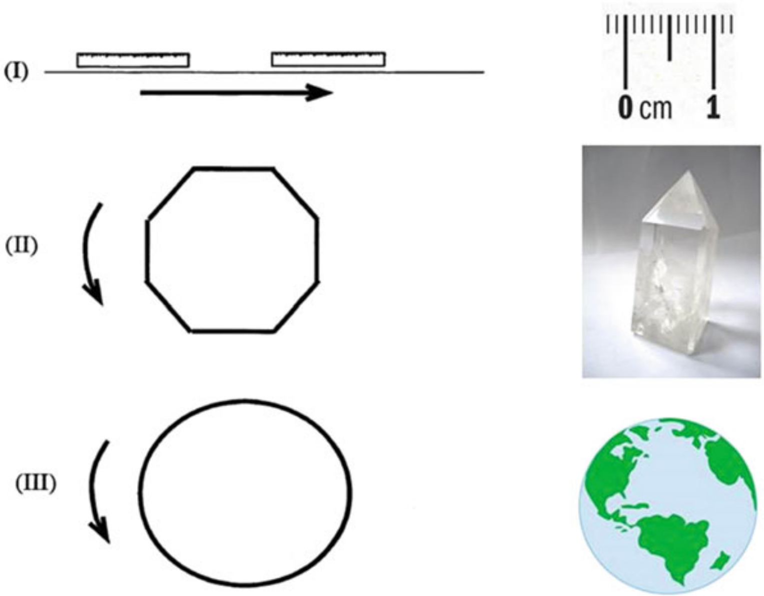

Fig. 3.4 Transformation symmetry and measurement units

task (in the Knox cube test (Sect. 5.3) or counting dots (Fig. 4.9)) where there is an obvious connexion: a more ordered task is easier to perform.

Secondly the principle of maximum entropy (most disorder) can be exploited when treating measurement error and uncertainty—so to say, the 'other side of the coin' to seeking the most ordered units. The loss of information on transmission through a measurement system can be modelled in terms of entropy (Eq. (3.3) in Sect. 5.4.1). Entropy can be in fact employed when characterising every element of a measurement system, for instance the commensurate ability of a person (or other probe) to perform the task. A more ordered person (less entropy) will be more able to perform, for instance the Knox block test[4] (Sect. 5.3).

In any process—for instance, transmission of measurement information through a measurement system—the change in entropy among all potential processes will be either zero or will increase as an example of the inexorable increase in disorder expressed by the second law of thermodynamics. Expressed mathematically:

$$H(Q|P) = H(P,Q) - H(P) \tag{3.3}$$

the entropy in the response, Q, of the measurement system given (conditional on) the stimulus input, P, is equal to the joint entropy of the message before and after

[4]This assumes of course that the necessary separation of object and instrument can be performed for the measurement system at hand (Sect. 3.1.2).

transmission minus the entropy of the original message. The joint entropy expresses how much information is transmitted and how much is lost or distorted (measurement uncertainty or bias) in the measurement process. The conditional entropy $H(Q|P)$ will be related (in Chap. 4, Eq. (4.5)) to a measure of the ability to perceive a dissimilarity in terms of a subjective distance between the distributions prior (P) and posterior (Q) to the measurement-based decision.

As explained by Klir and Folger (1988), the principle of maximum entropy applies when considering inferences whose content is beyond the evidence on hand. Otherwise called ampliative reasoning, one should 'use all but no more information than is available', that is, one should recognise and respect what one does not know—'knowing ignorance is strength' (Tsu 1972). Further discussion of measurement uncertainty will be found in Chap. 4 onwards.

3.3 Calibration, Accuracy and True Values

Measurement trueness goes beyond simply having small errors with respect to a local, perhaps arbitrary, reference value. Its essential meaning is in providing comparability of measurement results under both repeatability and reproducibility conditions—that is, measurements made by perhaps different laboratories using different equipment and operators. As society has developed and become more global, the need for measurement reference values of global applicability has increased in corresponding measure and a clear progression in metrological traceability to more universal measures is clear.

3.3.1 Trueness and Calibration Hierarchy

What defines a 'true' value, that is, the bull's eye of Fig. 2.2? Many years of discussion about the concept of measurement uncertainty (which is closely related but fundamentally different from metrological traceability (de Bièvre 2003)) has emphasised that—since true values are 'unknowable' and 'by nature indeterminate' quantities (to quote ISO GUM 1993 and VIM 2012, respectively)—then their use in measurement vocabularies is to be deprecated.

But in the context of metrological traceability there is one sense in which there is indeed access to what can be considered a true value—in fact in the very process of providing traceability, namely when making a calibration, that is, a measurement of a standard (or etalon) (Sect. 3.1.2). The value of a metrological standard (etalon), as determined previously with an uncertainty lower than in the actual measurements at hand, can be regarded as true—for instance, no one would consider the International Kilogram at the BIPM in Sèvres as giving anything but a true value of mass in any everyday weighing. The majority of measurements are made at lower levels in the

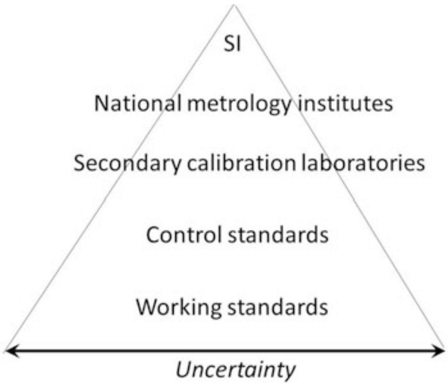

Fig. 3.5 Calibration hierarchy

calibration hierarchy[5] (Fig. 3.5) graded in order of measurement uncertainty and measures obtained from higher 'echelons' to all intents and purposes can be considered as true, since uncertainties in the standard are often much smaller than the uncertainty in the measurements at hand.

3.3.2 Objectivity and Calibration of Instruments in the Social Sciences

The process of calibration, i.e. tracing the measurement standards which, so to speak, embody defined amounts of the measurement unit, needs to be described in the case of a measurement system, including for the social sciences even systems where a human being is the instrument or other qualitative measurement situations.

In this book we identify two major factors which enable metrological traceability even in 'qualitative' situations such as found in the social sciences: (1) formulating measures of order in terms of entropy (Sect. 3.2.3) and (2) separating object and instrument (Sect. 3.1.2).

The Rasch model, mentioned in Sect. 1.4.5 and Eq. (1.3), is set up by performing a logistic regression to a set of data consisting of the responses of a number of instruments (e.g. people in a cohort) to a series of items (such as tasks).

A description of the calibration process where Man acts as measurement instrument in comparison with calibration of more traditional measurement systems in the physical sciences is given in Sect. 3.6.2.

While not enjoying access to universal units of measurement as in physics, in the social sciences—as in certified reference materials in analytical chemistry and materials science—one can establish recipes to define measurement units.

[5]Calibration hierarchy: 'sequence of calibrations from a reference to the final measuring system, where the outcome of each calibration depends on the outcome of the previous calibration' [VIM §2.40].

A construct specification equation (mentioned in the 'product' description Sect. 1.4.3 (Eq. (1.2))) will provide such a 'recipe' when formulating certified reference materials for traceability in the social sciences. A case study of the perception of the 'prestige' experienced with pre-packaged coffee will be evaluated in Sect. 4.5.2 as an example of the formulation of construct specification equations.

3.4 Politics and Philosophy of Metrology

3.4.1 Objective Measurement in the Physical and Engineering Sciences

The 'delicate' instruments of the physicist, referred to in *The Telegraphic Journal* 1884 (quoted in Gooday 1995), were not only used merely to make precise measurements (such as of the small electric currents in earlier telegraphy). The physicists' instruments also provided above all an 'absolute' accuracy, in other words a 'trueness', by which electrical quantities could be derived from the units of length, mass and time, the fundamental 'base' units of the metric system at that time. The universality and trueness of the latter were based on the ultimate physical reference of the era, namely the size and period of rotation of the Earth in true revolutionary universality '*A tous temps: A tous peuples*'. It took many years and was not until the electron was discovered at the turn of the nineteenth century before direct electrical measurements, with the voltmeter and ammeter, became to be raised in dignity and gain recognition as part of fundamental physics (Pendrill 2005, 2006a).

The same holds today in metrology: it is important to make precise measurements, in terms of low scatter or small uncertainty, as may be achieved by engineering a better measurement instrument. But perhaps arguably the main realm of the physicist in metrology is to provide for measurements which are traceable to absolute measures (ultimately, the universal fundamental constants). This enables the results of measurement to be related, not only of a particular quantity made by different people at different times and places (so important for trade and industry) but also to express different—apparently unrelated—quantities to each other in a more global sense. This latter universality of fundamental metrology relies on our understanding of the structure of the universe—spanning the realms of cosmology to elementary particle physics (Barrow 2002).

3.4.2 Politics and Trueness

In parallel with, and in some ways because of, the introduction of so-called neoliberal politics, views about how metrology was to be provided changed from about the 1980s, and this process is in a sense continuing for the foreseeable future.

A traditional view of a calibration hierarchy (Fig. 3.5) was where a national metrology institute (NMI) in each country (or perhaps the BIPM at the international level) would monopolise the top position at the pinnacle of the traceability pyramid for each measurement quantity. A more neoliberal approach would instead emphasise measurement laboratories of similar performance comparing their measurements, at a certain 'level' in the calibration hierarchy, such as a group of national metrology institutes aiming to refine the SI system for a particular measurement quantity. Previously exclusive consultative committees of the Metre Convention, had been the reserve of a few 'primary' NMIs charged with the tasking of recommending to the CIPM how each SI was to be defined. In the more neoliberal times, each consultative committee rapidly became open to many NMIs, albeit with some entry requirements demanding evidence of research.

'Customers', such as industrial metrology laboratories, would seek the 'best offer' from among a range of the NMIs as 'market actors'—a point of view which led by the turn of the Century to the Mutual Recognition Arrangement (MRA 1999) of the Metre Convention.

Another aspect of this neoliberal change in metrology was the increased emphasis on measurement uncertainty, to some extent at the expense of the more traditional fundamental concept of 'trueness'. Metrology was no longer to be 'only' furnishing unique true values, but instead one would allow for different, competing offers from whoever could convince the 'market'.

Over the years, NMIs have, as other organisations, been increasingly operated according to the neoliberal 'new public management (NPM) ' style, replacing former government 'authorities' which were perceived as inefficient and out-of-date. At the same, the private sector—which had pioneered this approach—increasingly found that NPM and neoliberalism were not optimal either...

3.4.3 Measurement Comparability in Conformity Assessment

Throughout the metrology vocabulary VIM, there is no mention of 'requirements' such as might be stipulated in conformity assessment, (apart from a few exceptions, such as 'maximum permissible measurement error' [VIM §4.26]). What might be called this 'general' perspective of measurement and quantities of the international metrology vocabulary is clear for instance in the Concept diagram for part of Clause 1 around 'quantity' of the VIM, where no explicit reference is made to either an entity or its quality characteristic. As mentioned in Sect. 2.2.2, one established area where conformity assessment of measurement instruments is regularly performed is legal metrology (Källgren et al. 2003; Pendrill 2014).

As explained in Chap. 1, the terminology of conformity assessment, in contrast to the metrological vocabulary, does emphasise a clear distinction (Fig. 1.3) between the quality characteristic $\eta = Z$ in the 'entity (or product) space' (Sect. 1.4.1) and the measurand of the quantity $\xi = Y$ in the 'measurement space' (Sect. 2.2.2). Metrological traceability provides the means of making measurement results comparable.

That measurement comparability is considered necessary to achieve the corresponding comparability of quality characteristics of entities, to the extent that such comparability is a requirement in conformity assessment. Comparability of entity quantities (or even kinds of quantity) is not always the prime concern in conformity assessment, where other factors—such as setting local safety limits—might be more pressing. In addition to the measurement value and any error in that estimate, it is also of interest in conformity assessment to deal with the variability—both real (in the entity) and apparent (as in measurement uncertainty). Uncertainty is not only expressed as a standard deviation, but in particular leads to risks of incorrect decisions of conformity. A pragmatic approach might be to set a limit on the measurement uncertainty, U_{cal}, associated with an uncorrected bias in relation to the tolerance maximum permissible error (*MPE*) in the context of testing for conformity assessment; perhaps stipulating that calibration uncertainties should not exceed half of the total measurement uncertainty, as an appropriate quantitative limit for when metrological traceability is significant in testing and measurement (Pendrill 2005, Sect. 4.3.2).

3.4.4 Objective Measurement in the Social Sciences

Physics has long been regarded as the original model for what a science should be. It has been a cherished hope and expectation of researchers in other disciplines that—given enough time, resources and talent—one should be able to achieve the same type of deep, broad and accurate knowledge achieved in the physical sciences, also in the biosciences or behavioural and social sciences. Research policy discussions often assume that all good science should be physics-like, i.e. characterized by the same quantitative specification of phenomena; a combination of mathematical sharpness and the deductive power of theory, and above all, a precise and profound understanding of the causes.

As Nelson (2015) has written: ‘Whereas physics can limit the subject matter it addresses so that such heterogeneity is irrelevant to its aims, for other sciences, this diversity or variability is the essence of what they study.’ As mentioned in Sect. 1.2.1, Nelson (2015) considers domains as diverse as cancer treatment, industrial innovation, K-12 education and environmental protection. Measurement uncertainty—as a measure of variability—is one the major hallmarks of metrology. We will later in this book give an account of a treatment of heterogeneity and the use of the concept of entropy when dealing with it.

Metrological references need to be founded on objective and sound measurement. Lacking an independent objective reality, e.g. in the social sciences, might lead to measurements providing no unique ‘right’ answer. This would make metrology of such qualitative assessments challenging, since: (1) independent reference standards used in metrology to ensure the comparability of different measurements would be difficult to establish separately from the actual measurement process, and

(2) measurement uncertainty would often be very large, since each new measurement set-up would produce definitions divergent from others.

In considering the philosophical foundations of social measurement, Maul et al. (2016) recall various approaches, including empiricism, pragmatism and realism. The philosophical realism behind physical metrology assumes, as in physics, that there is an objective reality, which exists even when we do not perceive or have instruments to measure it. One might argue—which Mari et al. (2016) refer to in terms of the output of their evaluation process—that there is 'seldom objective reality' in what is measured in social science (e.g. the challenge of a task) without our actually perceiving or measuring it. Mari et al. (2016) claim that a subjective opinion such as 'I am thirty percent happier today than I was yesterday' does not 'appear to deserve the trust that commonly accompanies measurement'.

Maul (2017), taking a broad perspective, argues that traditional approaches to the design and validation of survey-based measures may 'suffer from a number of serious shortcomings'. These include 'deeper confusions' regarding the relationship between psychological theory, modes of assessment and strategies for data analysis. Maul (2017) claims that operationalism may have encouraged the perception that psychological attributes need not be rigorously defined independently of a particular set of testing operations. He even states bluntly that the belief that measurement is a universally necessary component of scientific inquiry is 'simply not true'.

From our point of view, we would provide the counterargument that, for instance, opinions about fine art—e.g. the Mona Lisa painting by da Vinci—appear to be rather constant over the centuries and across different cultures. And measures of happiness are becoming essential components of person-centred care. A Rasch approach to perceived beauty or perceived happiness (or other pleasing patterns and degrees of order or symmetry) would in fact provide separate measures of the (albeit noisy) individual preferences of different persons and the intrinsic ability of, respectively, Leonardo's painting to stimulate pleasure or a particular activity of daily living to invoke happiness. This objectivity is perhaps not as strong as evaluations about the physical world (which would exist of course even without a human presence (Denbigh and Denbigh 1985)[6]), but is so to say 'fit for purpose' in the human-based context relevant for the present study. To use the vocabulary of the social sciences, such 'fit for purpose' references provide not only objectivity but—importantly—also intersubjectivity (Gillespie and Cornish 2010).

Our approach as presented in Sect. 3.3.2 can also be attributed to operationalism (as part of empiricism); i.e. defining a set of empirical operations performed with the measurement system. In that context, we circumvent the objections of realism since operationally it is 'meaningless to ask whether something is 'really' being measured' (Maul et al. 2016).

[6]'Definitional uncertainty'—'resulting from the finite amount of detail in the definition of a measurand' [VIM 2.27]—is of course in most cases much smaller in the strong objectivity of physics than in the social sciences.

In summary, the particular fusion of metrology and psychometrics proposed above, with its 'fit for purpose' objectivity and operationalism, appears to go some way in countering several of the philosophical reservations that had been expressed about attempting to quality-assure measurements in the social sciences (Pearce 2018).

3.5 Quantitative and Qualitative Scales

Table 3.3 summarises the different scales of Stevens, approximately compared with the data taxonomies suggested by Tukey. At the most elementary level, a categorical scale, called Nominal, is restricted to cases where no attempt is made to assign any order to the categories, which remain mere labels, being subordinate to all of the higher scales (ordinal, interval and ratio).

One step up from the nominal, towards a more quantitative scale is the case where, for intrinsic or extrinsic reasons, it is known that indications on a scale are generally ordered monotonically—i.e. a higher number indicates a higher measured quantity, although the exact, mathematical distance between marks on the scale is perhaps not known or investigated. Mathematically an ordinal scale can be expressed as:

$$x \geq y \text{ } iff \varnothing(\mathrm{x}) \geq \varnothing(\mathrm{y})$$

Such, so-called 'ordinal' scales are subordinate to the more familiar interval and ratio scales; that is, even the latter have ordinal properties but enjoy more quantitative properties. Responses on an ordinal scale can be assigned to a series of discrete categories, as illustrated in Fig. 1.1, although such discrete scales are of course not unique to ordinality and even ordinal scales can be depicted as continuous, as in the so-called visual analogue scales.

Table 3.3 Scales and data taxonomies

Scales of measurement (Stevens 1946)	Data taxonomies (Mosteller and Tukey 1977, Chap. 5)
Ratio	*Balances* (unbounded, positive or negative values)
	Amounts (non-negative real numbers)
Interval	*Counts* (non-negative integers)
	Counted fractions (bounded by zero and one. Includes percentages, e.g.)
Ordinal	*Ranks* (starting from 1, which may represent either the largest or smallest)
	Grades (ordered labels such as Freshman, Sophomore, Junior, Senior)
Nominal	*Names*

The challenges of treating measurement responses on an ordinal scale have been known since at least the late nineteenth century but surprisingly, well over a hundred years later, it is still common to find users uncritically applying the usual tools of statistics—e.g. calculating means, standard deviations, confidence intervals, multivariate analysis of variance—to indicated values on an ordinal scale where of course those tools cannot be certain to be valid if the underlying scale distances are not exactly known.

Particularly relevant when treating categorical measurement in both the physical and social sciences are counted fraction scales (Sect. 3.5.1), as distinct from other ordinal scales (Sect. 3.5.2). Common examples of 'counted fractions' are performance metrics for ability tests, customer satisfaction and decision risks caused by uncertainty. Kaltoft et al. (2014) emphasised the importance of rating overall decision quality, for instance in patient-centred care. Examples include binary test methods, where the response of the measurement system is a simple 'yes' or 'no', as for example in chemistry and microbiology where the responses mean 'species identified' or not. Other qualitative measures include examination of images and patterns, e.g. in analytical chemistry (Hardcastle 1998; Ellison and Fearn 2005), forensics (Vosk 2015), healthcare (Mencattini and Mari 2015) and more generally the performance of diverse systems where a 'probe' is used to investigate an 'item', such as the efficiency (Fig. 4.11).

At a step further towards a more quantitative scale lie counts, that is non-negative integers as the most elementary example of an interval scale which will be mentioned in the first case study in Sect. 3.6.1.

3.5.1 Counted Fractions

One common family of ordinal scale is termed 'counted fractions', i.e. relative-number problems famously summarised by Tukey (1986), such as counting for example relative fractions of 'how many sheep & goats; are affected at this dose; or pebbles are quartz', and so on. As early as 1897, Pearson warned of the dangers of counting fractions: 'Beware of attempts to interpret correlations between ratios whose numerators and denominators contain common parts'. Mathematically, in the counted fraction expression, $X_j\% = \frac{X_j}{\sum_i X_i}$, the presence of the amount X of component j appearing in both the numerator and denominator means—and increasingly where X_j is either large or small compared with the other components—that any error in X_j will be correlated with the other components, since of course there is the boundary condition $\sum_j X_j\% = 100\%$. In the days before computers in the first half of the twentieth century, such non-linearities at the scale extremities—both the high and low ends—although known to be linearisable through so-called logistic ruling, converting percentages p into logits, l, with the expression $l = \ln\left(\frac{p}{100-p}\right)$, were

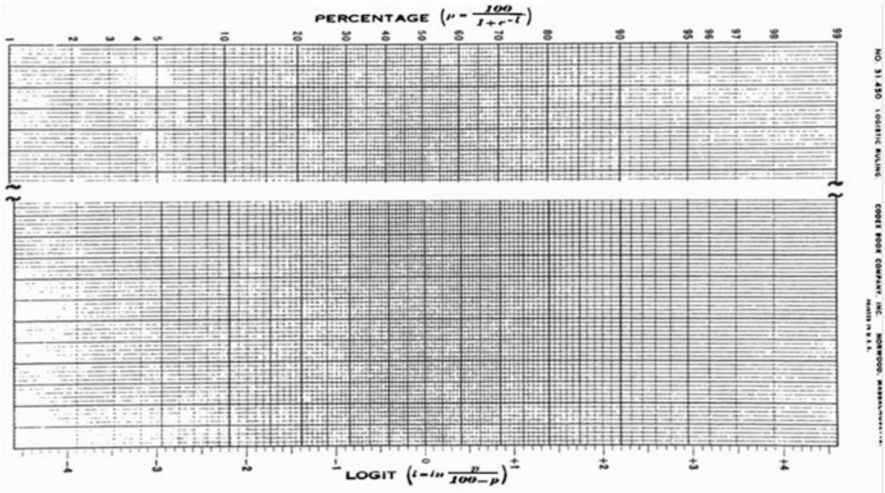

Fig. 3.6 Logistic ruling (Tukey 1986, reproduced with permission Taylor & Francis)

arduous to calculate, and it was common to use nomographs, such as exemplified in Fig. 3.6.

Throughout the century, the counted fraction dilemma has re-emerged in various schools. Aitchison (1982) for instance in what he termed compositional data analysis met with resistance in groups he termed:

Wishful Thinkers

No problem exists or, at worst, it is some esoteric mathematical statistical curiosity which has not worried our predecessors and so should not worry us. Let us continue to calculate and interpret correlations of raw components. After all if we omit one of the parts, the constant-sum constraint no longer applies. Someday, somehow, what we are doing will be shown by someone to have been correct all the time.

Describers

As long as we are just *describing* a compositional data set we can use *any* characteristics.
In describing compositional data we can use:

- arithmetic means,
- covariance matrices of raw components
- indeed any linear methods
- such as principal components of the raw components.

After all we are simply describing the data set in summary form, not analyzing it.

3.5.2 Other Ordinal Scales. Pragmatism

While common, the counted fraction scale (Sect. 3.5.1) is not the only example of an ordinal scale, and it is important to bear in mind that the logistic transformation used in the Rasch approach does not automatically fix all ordinality problems.

In Fig. 3.7 is shown a type of ordinal scale where the measured depth in metres of various geological strata do not of course directly correspond to the meaningful scales of deposition over different epochs and geological processes.

This and similar kinds of initially unknown scale may well be present together with the basic counted fraction ordinality, thus requiring special diagnostic tools to detect scale non-linearities and multidimensionality. Generalized linear models (GLM) (McCullagh 1980) can handle cases where the response variable (*R*) cannot always be expected to vary linearly with the explanatory variable (*S*).

Other examples are pragmatic scales (Table 3.1) where a cost function can be employed to introduce a distance metric on an otherwise weakly defined ordinal scale (Bashkansky et al. 2007), thus going beyond the traditional limitations of statistical measures of location and dispersion on such scales as well as of course capturing the impact of particular measurement 'values' in the broadest sense (Weinberger 2003). Examples may be found in legal metrology where commodities (such as petrol or pre-packaged goods) are normally priced linearly by quantity (Pendrill 2014). The impact of customer dissatisfaction on the other hand, for instance, with short measures of a commodity may well depend quadratically with increasing discrepancies, ε, of quantity from the expected value, as commonly expressed by the Taguchi loss function Cost $\cdot (k \cdot \varepsilon)^2$, according to the impact of incorrect classification where 'Cost' is the cost per unit squared assignment error, ε^2, with respect to the expectation (or nominal value) (Pendrill 2006b). Such pragmatic

Fig. 3.7 Ordinal scale in geology. Photo: courtesy of Florence Pendrill

measures will be employed in the final chapter of this book when making decisions about product based on measurement.

3.6 New and Future Measurement Units

3.6.1 The Revised SI

There is much written elsewhere about the revised SI, for instance, as referenced on the Metre Convention website (BIPM 2018), to which we refer the reader for a more comprehensive account.

Classic examples of the universality of measurements in the framework of the laws of physics are the estimations of fundamental physical constants, such as the elementary electronic charge or the Boltzmann constant (Pendrill 1994). Completely different physical experiments which at first appearance are apparently unrelated—such as measurements of voltage in a semiconductor at cryogenic temperatures and a measurement made of the magnetic moment of a single electron in an electrostatic trap—can nevertheless yield consistent estimates of a particular fundamental physical constant. Traceability in the framework of the laws of physics thus enables apparently unrelated measurement quantities to be compared (Quinn 1994). The least-squares adjustment of the values of fundamental physical constants, as made periodically by the CODATA Task Force, is based on this (CODATA 2004).

De Boer (1994/5) writes: 'It is important to note that Maxwell writes that the *unit has also to be conceived as a quantity…*'. This statement does not need reassessment even though, as in the revised SI (CGPM 2018, 2019), the majority of SI units are defined in terms of fundamental constants. A careful choice of terminology would be, for example: 'the speed of light in vacuum *c has the value* is 299 792 458 m/s, …' (adapted from CGPM 2018).

Table 3.1 indicates how measurement units are expressed depending on which level in the hierarchy of concepts in quantity calculus is most relevant and meaningful to the task at hand.

Clear Definitions

A correct terminological formulation of measurement units in the SI system is particularly important, whose very role is to communicate measurement information meaningfully and widely; as summarised by Petley (1990) to everyone 'from the Nobel Prize winner to the proverbial man and woman in the street'. In the revised SI from 2018, a distinction is made between the new 'explicit constant'[7] and a more 'explicit unit' traditional definitions of units. A challenge posed by the explicit constant definitions of the revised SI is that they might be perceived as somewhat more abstract than the more traditional explicit unit definitions for some readers not familiar with fundamental physical constants.

Reasons for this can be that the revised definitions are formulated (1) in terms of the least informative concept level in quantity calculus, namely as 'quantity values' opposed to 'quantities' (Mari et al. 2018), and (2) assume a certain pre-requisite knowledge: In a table of concepts expected to be included in the terminology of a unit definition, several cells in the explicit constant definition table are left blank (–):

Explicit constant definition: unit of length

Unit	Kind of quantity	Quantity	Expression, equation of physics	Unit entity[a]	Values of constants	Related quantities	Related units
Metre (m)	Length	–	–	–	$c = 299\,792\,458\ m/s$	Time (t)	Second

[a]Maxwell's 'individual thing' (de Boer 1994/5)

These cells become filled in the more readily understandable explicit unit definition table:

Explicit unit definitions: Unit of length

Unit	Kind of quantity	Quantity[a]	Expression, equation of physics	Unit entity	Values of constants	Related quantities	Related units
Metre (m)	Length	Pathlength (ℓ)	$\ell = c \cdot t$	Path of light in vacuum	$c = 299\,792\,458\ m/s$	Time (t)	Second
Metre (m)	Length	Wavelength (λ)	$\lambda = \frac{c}{\nu}$	Wavelength of light in vacuum	$c = 299\,792\,458\ m/s$	Frequency (ν)	1/s

[a]According to quantity calculus, only quantities of the same kind have the same unit

In motivating the new approach of explicit constant definitions of the revised SI it is claimed that: 'A user is now free to choose any convenient equation of physics that links the defining constants to the quantity intended to be measured' (SI Brochure, 9th edition, 2.3.2). Two examples of such expressions are given in the explicit unit definition table above. The freedom of choice and the increased accuracy of fundamental physical constants offered by explicit constant definitions appears however to have to be traded against an increased difficulty—particularly for those unfamiliar with fundamental physics—in conveying meaning and understanding.

[7]'That is, a definition in which the unit is defined indirectly by specifying explicitly an exact value for a well-recognized fundamental constant' [24th CGPM, 2011 On the possible future revision of the International System of Units, the SI (CR, 532), Resolution 1].

Quantum Mechanics and Measurement

Dirac (1992) makes a brief mention of the special case when the real dynamical variable is a number, every state is an eigenstate and the dynamical variable is obviously an observable.

Any measurement of it always gives the same result, so it is just a physical constant, like the charge on an electron. A physical constant in quantum mechanics may thus be looked upon:

- either as an observable with a single eigenvalue
- or as a mere number appearing in the equations,

the two points of view, according to Dirac, being equivalent.

In line with our discussion of measurement units in the context of symmetry and entropy (Sect. 3.2.3), we would like to go beyond regarding a physical constant as a 'mere number'. One can attempt to describe the amount of information in terms of a sum of 'chunks' of the information in a message (Sect. 3.2.3), that is one seeks with a construct specification equation (Stenner et al. 1983, 2013; Sect. 1.4.3 (Eq. (1.2))) to resolve the information in the message into a number of explanatory variables which are the 'irreducible representations' corresponding to the eigenvectors, such as a series of spherical harmonics used to describe an arbitrary signal.

A connexion can be readily made between quantity calculus (Sect. 3.1) and the matrix formulation of quantum mechanics and the corresponding formulation in classical mechanics.[8]

Measurement is frequently mentioned when describing quantum mechanics, often referring to the impossibility—because of the non-zero value of the Planck constant—of making a measurement without disturbing the measurement object. This famously includes the Heisenberg uncertainty relation and contemporary topics in physics such as quantum entanglement and the possibilities of making quantum computers.

In the present context, we would like to highlight at the basic level how the measurement process is described in quantum mechanics, as a template for a description of measurement more broadly, in line with the aim of this book to give a unified view of measurement in the physical and social sciences. Two aspects in particular will be highlighted: (1) the distinct concepts of measurement object, quantity and quantity value and (2) measurement as a displacement process.

When discussing a general physical interpretation of a mathematical theory of quantum mechanics, Dirac (1992, §12, p. 45) recalls that in 'classical mechanics an observable always "has a value" for any particular state of the system' and clearly mentions measurement aspects: 'When we make an observation we measure some

[8]By the correspondence principle, relations specific to quantum mechanical effects, e.g. on the microscopic scale, can find correspondence to relations in Newtonian physics, e.g. at the macroscopic scale where the Planck constant is negligibly small.

dynamical variable. It is obvious physically that the result of such a measurement must always be a real number,...'

Eigenstates play of course a well-known role in quantum mechanics, and when introducing the expression:

$$\mathbf{Q}|q\rangle = q|q\rangle \tag{3.4}$$

Dirac (1992, §10, p. 35) emphasises the measurement aspects: 'If the dynamical system is in an eigenstate of a real, dynamical variable **Q**, belonging to the eigenvalue q, then a measurement of **Q** will certainly give as result the number q....'. Extension of this formulation to include functions of observables is described in §11 Dirac's (1992) book.

Dirac points out (1992, §12, p. 46) that one can 'extract physical information from the mathematics even when we are not dealing with eigenstates' provided one assumes that, 'if the measurement of the observable **Q** for the system in the state corresponding to $|q\rangle$ is made a large number of times, the average of all the results obtained will be $\langle q|\mathbf{Q}|q\rangle$, provided $|q\rangle$ is normalised.'

Interestingly, Dirac's formula (3.4) clearly includes explicit reference to the *entity* being measured by including $|q\rangle$. One can compare this with our description of measurement systems (Chap. 2). Note that the Dirac formulation also clarifies that **Q** on the LHS of (3.4) denotes a *quantity*—namely, the measurand associated with the measurement object (entity), as distinct from a *quantity value* (which would be the eigenvalue, q, on the RHS of (3.4)). This can be set in relation to our discussion of the hierarchy of measurement concepts in quantity calculus (Sect. 3.1.1 and Table 3.2).

A second aspect we would like to highlight is how *measurement as a displacement process* enters into a description of quantum mechanics. Recall Maxwell's classical words in connexion with Eq. (3.2) of 'making up' a required quantity, and counting how many times a unit fits in the measured displacement, where 'displacement' is not specifically in length, but in the dimension of interest. The aim of the displacement can be interpreted in the broadest sense, not just in the physicist's laboratory, but more generally including the measurement comparability needed in any application (Sect. 3.2).

Dirac (1992, §25) writes about displacement operators which can be seen as a description of a measurement process: 'The displacement of a state or observable is a perfectly definite process physically. Thus to displace a state or observable through a distance δx in the direction of the x-axis, we should merely have to displace all the apparatus used in preparing the state, or all the apparatus required to measure the observable, through the distance δx in the direction of the x-axis, and the displaced apparatus would define the displaced state or observable. ... A displaced state or dynamical variable is uniquely determined by the undisplaced state or dynamical variable together with the direction and magnitude of the displacement'.

Dirac (1992, §26) goes on to describe a displacement operator, D, such that for any dynamical variable v, a displaced dynamical variable $v_d = DvD^{-1}$. One particular kind of displacement operator is a linear *unitary* transformation operator U. This special displacement operator can transform any linear operator Q to a corresponding linear operator $Q*$ according to the relation: $Q^* = U \cdot Q \cdot U^{-1}$ so that each Q∗ has the *same* eigenvalues as the corresponding Q, $Q|q\rangle = q|q\rangle$ and hence: $Q^* U|q\rangle = qU|q\rangle$.

A connexion between Dirac's and Maxwell's descriptions of displaced measurement systems can be established by observing that the unit quantity has an eigenvalue q_{unit}: $[Q]|q\rangle = \{q_{\text{unit}}\}|q\rangle$, and a 'displaced' observable: $Q^*|q\rangle = \{Q\} \cdot [Q]|q\rangle = \{Q\} \cdot \{q_{\text{unit}}\}|q\rangle$, by simply combining (3.4) and (3.2), so that:

$$U = \{q_{\text{unit}}\} \cdot [Q] \tag{3.5}$$

The quantisation rule: $\oint p \cdot dq = n \cdot h$ for a pair of canonical variables (p, q) such as position and momentum or time and energy (Born 1972), means that, in one period of the motion, the integral yields an area which is an integral multiple of h, according to the quantum postulate. The eigenfunctions of orbital angular momentum operator, as irreducible representations, form a set of base functions accompanied by a set of eigenvalues (multiples of h) which can be regarded as measurement units (or 'chunks') (Sect. 3.2.3). For our purposes in metrology, it illustrates how the Planck constant h acts as a fundamental measurement unit.[9]

How the fundamental physical constants now enter into definitions of measurement units in the revised SI is described further in 9th ed SI Brochure (CGPM 2019, CGPM 2018, 2019). In the present description, one can seek units of measurement more generally from fundamental symmetries described in terms of minimum entropy (Sect. 3.2.3).

The examples given in the remaining sections of this chapter in different ways have to do with counting.

Boltzmann Constant and Elementary Counting

Atomic helium and its microscopic structure have been the focus of attention of a number of highly accurate experimental and theoretical investigations of properties such as electronic binding energies, fine structure, etc., which are leading in some cases to new estimates of fundamental constants, such as α, the fine structure constant. At the same time, the thermodynamic properties of *macroscopic* helium gas are being studied intensively, particularly in connection with the development of gas thermometry. Connecting these microscopic and macroscopic studies of helium,

[9] The Planck constant is not merely a 'number' but has multiplicity of roles, such as (1) a constant of proportionality between canonical pairs of quantities (e.g. energy/time, momentum/position, angular momentum/rotation); (2) acting as a fundamental unit as the 'quantum' of 'action' (e.g. energy·time); (3) is implicit in many of the 'quantum' definitions of the SI, not only the kilogram but also the second, volt and ohm; (4) quantifying the interaction through fields between physical systems, such as the electromagnetic interaction mediated by 'virtual' photons (Cohen-Tanoudji 1993).

Pendrill (1996) proposed a new estimate of the Boltzmann constant, k, as well as a general assessment of the reliability of optical and electrical measurements of the polarisation properties of helium gas. The link between the macroscopic (gas constant, R, from dielectric constant gas thermometry (DCGT) of ^{4}He) and the microscopic (Boltzmann constant) is provided by the relation: $R = N_A \cdot k$, where the Avogadro constant, N_A, represents counting the number of elementary microscopic entities to make up a macroscopic quantity. The 1996 estimate proposed that the Boltzmann constant, k had the value 1.380 628(16) $\times$ 10^{-23} J/K (provided one accepted the *ab initio* relativistic value of the atomic polarizability and the DCGT experimental results available at the time) can be compared with the most recent value proposed in the new SI: 1.380 649 $\times$ 10^{-23} J/K (CGPM 2018, 2019).

Counts and Quantities of Unit One

An example of elementary counting (of the number of dots) regarded as measurement is given in Fig. 4.9.

Whether 'counts' (non-negative integers), of for example the number of pills, can be considered as 'dimensionless quantities in the SI' is still an active subject of debate in the international literature (Flater 2017). As remarked by Kogan (2014), a 'quantity for which all exponents of factors corresponding to base quantities in its quantity dimension are zero', is preferably called a quantity of unit one (where 'dimension' is discussed in Sect. 3.2.2).

The introduction of the number of entities as a base quantity can help answer the question: 'is amount of information a quantity?' (Kogan 2014) The amount of information usually stands for measure of information in a message. This quantity has its own unit in the informatics—it is called bit and defined as a unit of the amount of information in the binary computing, the minimal unit of the amount of information that can be transmitted or stored, and corresponds to one binary digit that can have one of the two values, 0 or 1. One of the earliest and elementary measures of information was formulated by Hartley (Klir and Folger 1988) in terms of how specific a particular sign is:

$$I(N) = \log_b(N) \tag{3.6}$$

where N is the number of distinguishable signs. The least amount (a 'quantum') of information—one bit—corresponds to $N = 2$ signs (e.g. 0 & 1) on a base of $b = 2$. Equation (3.6) can be converted into terms of informational entropy by multiplying by the Boltzmann constant, k, which, together with the bit, play a role in 'information quantisation' analogous to that of the Planck constant in quantum mechanics (Cohen-Tannoudji 1993). We return to the connexion between entropy and information in our studies of cognitive tests in neuropsychology and elementary counting (Chap. 4).

3.6.2 Human Challenges

In its logistic regression form, the 'straight ruler' aspect of the Rasch formula, i.e. Eq. (1.1), has been described by Linacre and Wright (1989) in the following terms: 'The mathematical unit of Rasch measurement, the log-odds unit or "logit", is defined prior to the experiment. All logits are the same length with respect to this change in the odds of observing the indicative event.'

The Rasch invariant measure approach goes further in defining measurement units (Humphry 2011) since it uniquely yields 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', i.e. specific objectivity (Irwin 2007). The Rasch (1961) approach is not simply mathematical or statistical, but instead a specifically metrological approach to human-based measurement. Note that the same probability of success can be obtained with an able person performing a difficult task as with a less able person tackling an easier task. The separation of attributes of the measured item from those of the person measuring them brings invariant measurement theory to psychometrics. Fisher's (1997) work on the metrology of instruments for physical disability (Fig. 3.8) was one of the first to demonstrate the concepts of intercomparability through common units, where item banking is a common expression (Pesudovs 2010).

Having enabled with Rasch Measurement Theory a set of metrological references, e.g. for task difficulty, one can then proceed to set up a scale (analogous to conventional measurement etalons (Fig. 3.9)) which is delineated by measurement units where any measured quantity, $\delta_j = \{\delta_j\} \cdot [\delta]$, is the product of a number { } and a unit denoted in square brackets [ ]—Eq. (3.2). This step is enabled by combining a

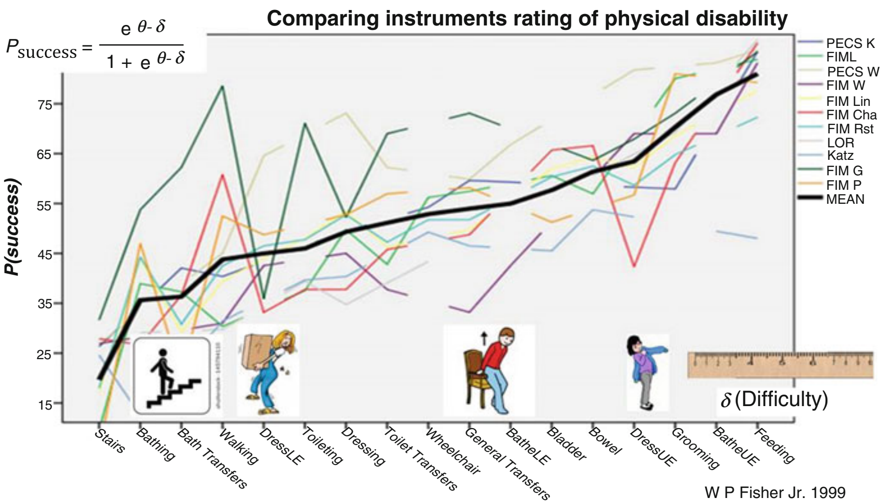

Fig. 3.8 Probability of succeeding for a range of physical disability tasks of different difficulty (courtesy: W P Fisher Jr.)

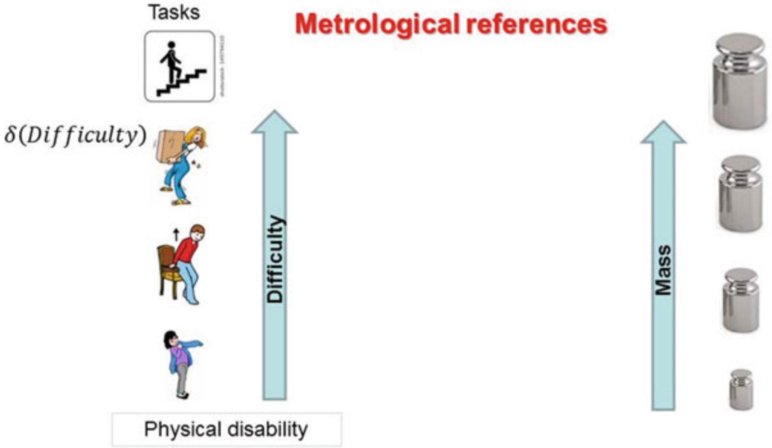

Fig. 3.9 A set of tasks of increasing difficulty as metrological references analogous to a set of mass standards

procedure to transform qualitative data to a new 'space' (in the present case, through restitution, to the space of the measurand, as illustrated in Figs. 2.11 and 3.3), together with ability of Rasch Measurement Theory to provide separate estimates of measurement and object dispersions in the results when Man acts as a measurement instrument.

This new approach to the metrological treatment of qualitative data differs from others in that the special character of the qualitative data is assigned principally not to the measurand but to the response of the measurement system. (One can draw analogies with the common expression: 'Beauty is in the eye of the beholder'.) Using Rasch Measurement Theory in the restitution process re-establishes a linear, quantitative scale for the measurand (e.g. for a property such as task difficulty) where metrological quality assurance—in terms of traceability and uncertainty—can be performed.

Measurement Units and the Rasch Model

At the mathematical level, it is not immediately obvious how measurement units can be inserted in the Rasch model, which is a logarithmic function. In order to include measurement units explicitly in general linearised models, such as the Rasch model, Humphry (2011) proposed a modified version of Eq. (1.1), called a 'logistic measurement function':

$$\ln\left(\frac{P_{\text{success},i,j}}{1 - P_{\text{success},i,j}}\right) = \rho_s \cdot \left(\theta_i^* - \delta_j^*\right) \tag{3.7}$$

where s indicates a classification of an empirical factor; ρ is a multiplicative constant; and the modified Rasch parameters are related to the original Rasch parameters through the expressions $\theta_i^* = \frac{\theta_i}{\rho}$ and $\delta_j^* = \frac{\delta_j}{\rho}$. If units are to be associated with person and item attributes, respectively, as $\theta_i = \{\theta_i\} \cdot [\theta]$ and $\delta_j = \{\delta_j\} \cdot [\delta]$ then assuming that item and person attributes share the same scale—a key aspect of the Rasch model—gives an expression for the 'common unit' of measure as: $[\theta] = [\delta]$ (denoted $[u_*]$ by Humphry (2011)).

Equation (3.7) appears on first sight to be similar to Item Response Theory expressions, but there is a subtle distinction, as expressed recently: 'Item Response Theory models are statistical models used to explain data, and the aim of an Item Response Theory analysis is to find the statistical model that best explains the observed data. By contrast, the aim of Rasch Measurement Theory is to determine the extent to which observed clinical outcome assessment data satisfy the measurement model' (Barbic and Cano 2016).

As pointed out by Humphry and Andrich (2008), the incorporation in an Item Response Theory model of a discrimination parameter which is estimated for each item (or person) will in general break conditions for sufficiency and specific objectivity, and thus the opportunity of establishing units and measurement scales. But this opportunity is maintained if one, as in Eq. (3.7), associates a discrimination factor (ρ) with a *set* of items rather than a single item, according to Humphry and co-workers (Humphry 2011; Asril and Marais 2011), as will be exemplified in Sect. 5.4.1.

The concept of entropy will be a main guide in this book when formulating construct specification equations, be it of task difficulty or person (instrument) ability (Sect. 5.1). The increasing difficulty of remembering block sequences (such as the Corsi block test) was described by Schnore and Partington (1967) in terms of a sum of a set of basic patterns (or chunks) with different information content expressed in terms of entropy and the symmetry of each pattern. These can be regarded as measurement units in a similar fashion to the irreducible representations used to describe arbitrary functions such as in quantum mechanics (Sects. 3.2.3 and 3.6.1).

References

J. Aitchison, The statistical analysis of compositional data. J. R. Stat. Soc. **44**, 139–177 (1982)

A. Asril and I. Marais, (2011) Applying a Rasch Model Distractor Analysis. In: Cavanagh R.F., Waugh R.F. (eds) Applications of Rasch Measurement in Learning Environments Research. Advances in Learning Environments Research, vol **2**. SensePublishers, Rotterdam

S. Barbic and S. Cano, The application of Rasch measurement theory to psychiatric clinical outcomes research, BJPsych Bulletin, **40**, 243–244, https://doi.org/10.1192/pb.bp.115.052290 (2016)

F. Attneave, Informational aspects of visual perception. Psychol. Rev. **61**, 183–193 (1954)

H. Barlow, The exploitation of regularities in the environment by the brain. Behav. Brain Sci. **24**, 602–607 (2001)

J. Barrow, *From Alpha to Omega* (Jonathan Cape, London, 2002). ISBN 0224061356

E. Bashkansky, S. Dror, R. Ravid, P. Grabov, Effectiveness of a product quality classifier. Qual. Eng. **19**(3), 235–244 (2007)

J.P. Bentley, *Principles of Measurement Systems*, 4th edn. (Pearson, Prentice-Hall, Lebanon, 2004). ISBN-13: 978-0130430281, ISBN-10: 0130430285

BIPM, On the Future Revision of the SI, (2018), https://www.bipm.org/en/measurement-units/rev-si/

M. Born, *Atomic Physics*, 8th edn. (Blackie & Son Ltd., London, 1972). 216.89027.6, ISBN-13: 978-0486659848, ISBN-10: 0486659844

N.R. Campbell, *Physics – The Elements* (Cambridge University Press, Cambridge, 1920)

CGPM, SI Brochure: The International System of Units (SI), 9th edn. (2019), https://www.bipm.org/utils/common/pdf/si-brochure/SI-Brochure-9.pdf

CGPM, On the revision of the International System of Units (SI), in *Draft Resolution* **A** – 26th meeting of the CGPM (13–16 November 2018), (2018), https://www.bipm.org/utils/en/pdf/CGPM/Draft-Resolution-A-EN.pdf

CODATA, *Task Force on Fundamental Physical Constants, Committee on Data for Science and Technology* (ICSU, Paris, 2004). http://www.codata.org/taskgroups/TGfundconst/index.html

G. Cohen-Tannoudji, Universal Constants in Physics (MCGRAW HILL HORIZONS OF SCIENCE SERIES), ISBN-13: 978-0070116511, McGraw-Hill Ryerson, Limited (1993)

P. de Bièvre, Traceability is not meant to reduce uncertainty. Accred. Qual. Assur. **8**, 497 (2003)

J. de Boer, On the history of quantity calculus and the international system. Metrologia **31**, 405–429 (1994/5)

K.G. Denbigh, J.S. Denbigh, *Entropy in Relation to Incomplete Knowledge* (Cambridge University Press, Cambridge, 1985). ISBN 0 521 25677 1

P.A.M. Dirac, The principles of quantum mechanics, in *The International Series of Monographs on Physics*, ed. by J. Birman et al., 4th edn., (Clarendon Press, Oxford, 1992)

S. Ellison, T. Fearn, Characterising the performance of qualitative analytical methods: statistics and terminology. TRAC-Trend Anal. Chem. **24**, 468–476 (2005)

W.H. Emerson, On quantity calculus and units of measurement. Metrologia **45**, 134–138 (2008)

EN 15224:2012, *Health care services – Quality management systems – Requirements based on EN ISO 9001:2008*

W.P. Fisher, Jr., Physical disability construct convergence across instruments: Towards a universal metric. Journal of Outcome Measurement **1**(2), pp 87–113 (1997)

R. Feynman, The Feynman Lectures on Physics, vol. I, (2013), http://www.feynmanlectures.caltech.edu/I_01.html#Ch1-S1

D. Flater, Redressing grievances with the treatment of dimensionless quantities in SI. Measurement **109**, 105–110 (2017)

R. Fleischmann, Einheiteninvariante Gröβengleichungen, Dimension. Der Mathematische und Naturwissenschaftliche Unterricht **12**, 386–399 (1960)

A. Gillespie, F. Cornish, Intersubjectivity: towards a dialogical analysis. J. Theory Soc. Behav. **40**, 19–46 (2010)

G. Gooday, in *The Values of Precision*, ed. by M. N. Wise, (Princeton University Press, Princeton, 1995). ISBN 0-691-03759-0

W. Hardcastle, *Qualitative Analysis: A Guide to Best Practice* (Royal Society of Chemistry, Cambridge, 1998)

S.M. Humphry, The role of the unit in physics and psychometrics. Meas. Interdiscip. Res. Perspect. **9**(1), 1–24 (2011)

S.M. Humphry, D. Andrich, Understanding the unit implicit in the Rasch model. J. Appl. Meas. **9**, 249–264 (2008)

R.J. Irwin, A psychophysical interpretation of Rasch's psychometric principle of specific objectivity. Proc. Fechner Day **23**, 1–6 (2007)

JCGM200:2012 International vocabulary of metrology—basic and general concepts and associated terms (VIM 3rd edition) (JCGM 200:2008 with minor corrections) WG2 Joint Committee on Guides in Metrology (JCGM) (Sevrès: BIPM)

H. Källgren, M. Lauwaars, B. Magnusson, L.R. Pendrill, P. Taylor, Role of measurement uncertainty in conformity assessment in legal metrology and trade. Accred. Qual. Assur. **8**, 541–547 (2003)

M. Kaltoft, M. Cunich, G. Salkeld, J. Dowie, Assessing decision quality in patient-centred care requires a preference-sensitive measure. J. Health Serv. Res. Policy **19**, 110–117 (2014). https://doi.org/10.1177/1355819613511076

G.J. Klir, T.A. Folger, *Fuzzy sets, uncertainty and information* (Prentice Hall, New Jersey, 1988). ISBN 0-13-345984-5

J. Kogan, An Alternative Path to a New SI, Part 1: On Quantities with Dimension One, (2014), https://web.archive.org/web/20160912224601/http://metrologybytes.net/PapersUnpub/Kogan_2014.pdf

L.D. Landau, E.M. Lifshitz, Mechanics, Course of Theoretical Physics (Butterworth-Heinenann, 1976), 3rd ed., Vol. 1. ISBN 0-7506-2896-0.

B. Langefors, Essays on Infology, in *Gothenburg Studies of Information Systems*, ed. by B. Dahlbom. Report **5** (University of Göteborg, 1993)

J.M. Linacre, B. Wright, The 'length' of a Logit. Rasch Meas. Trans. **3**, 54–55 (1989)

L. Marco-Ruiz, A. Budrionis, K.Y. Yigzaw, J.G. Bellika, Interoperability Mechanisms of Clinical Decision Support Systems: A Systematic Review, in *Proceedings of the 14th Scandinavian Conference on Health Informatics*, Gothenburg, Sweden, April 6–7, 2016, (2016). http://www.ep.liu.se/ecp/122/ecp16122.pdf

L. Mari, A. Maul, D. Torres Irribarra, M. Wilson, A metastructural understanding of measurement. J. Phys. Conf. Ser **772**, 012009 (2016). IMEKO2016 TC1-TC7-TC13

L. Mari, C.D. Ehrlich, L.R. Pendrill, Measurement units as quantities of objects or values of quantities: a discussion. Metrologia **55**, 716 (2018). https://doi.org/10.1088/1681-7575/aad8d8

A. Maul, D. Torres Irribarra, M. Wilson, On the philosophical foundations of psychological measurement. Measurement **79**, 311–320 (2016)

A. Maul. Rethinking traditional methods of survey validation, Measurement. Interdisciplinary Research and Perspectives **15**(2), 51–56 (2017)

P McCullagh, Regression models for ordinal data. J. Roy. Stat. Soc., 42: p. 109–42 (1980)

A. Mencattini, L. Mari, A conceptual framework for concept definition in measurement: the case of 'sensitivity. Measurement **72**, 77–87 (2015)

G.A. Miller, The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**, 81–97 (1956)

F. Mosteller, J.W. Tukey, *Data Analysis and Regression: A Second Course in Statistics* (Addison-Wesley, Reading, 1977)

MRA, International equivalence of measurements: the CIPM MRA, (1999), https://www.bipm.org/en/cipm-mra/

R.R. Nelson, Physics envy: get over it. Iss. Sci. Technol. **XXXI**, 71–78 (2015). http://issues.org/31-3/physics-envy-get-over-it/

J. Pearce, Psychometrics in action, science as practice. Adv. Health Sci. Educ. **23**, 653–663 (2018). https://doi.org/10.1007/s10459-017-9789-7

L.R. Pendrill, Assuring measurement quality in person-centred healthcare. Meas. Sci. Technol **29**(3), 034003 (2018). https://doi.org/10.1088/1361-6501/aa9cd2. special issue Metrologie 2017.

L.R. Pendrill, Some comments on fundamental constants and units of measurement related to precise measurements with trapped charged particles. Phys. Scripta **T59**, 46–52 (1994). (1995) and World Scientific Publishing, ed. I Bergström, C Carlberg & R Schuch, ISBN 981-02-2481-8 (1996)

L.R. Pendrill, Macroscopic and microscopic polarisabilities of helium gas. J. Phys. B At. Mol. Opt. Phys. **29**, 3581–3586 (1996)

L.R. Pendrill, Meeting future needs for metrological traceability – a physicist's view. Accred. Qual. Assur. **10**, 133–139 (2005). http://www.springerlink.com/content/0dn6x90cmr8hq3v4/?p=2338bc01ade44a208a2d8fb148ecd37a&pi

L.R. Pendrill, Metrology: time for a new look at the physics of traceable measurement? Europhysics News **37**, 22–25 (2006a). https://doi.org/10.1051/epn:2006104

L.R. Pendrill, Optimised measurement uncertainty and decision-making when sampling by variables or by attribute. Measurement **39**(9), 829–840 (2006b). https://doi.org/10.1016/j.measurement.2006.04.014

L.R. Pendrill, Uncertainty & risks in decision-making in qualitative measurement, in *AMCTM 2011 International Conference on Advanced Mathematical and Computational Tools in Metrology and Testing*, Göteborg June 20–22 2011, (2011), http://www.sp.se/AMCTM2011
L.R. Pendrill, Using measurement uncertainty in decision-making & conformity assessment. Metrologia **51**, S206 (2014)
K. Pesudovs, Item banking: a generational change in patient-reported outcome measurement. Optom. Vis. Sci. **87**(4), 285–293 (2010). https://doi.org/10.1097/OPX.0b013e3181d408d7
B.W. Petley, *The fundamental physical constants and the frontier of measurement* (Adam Hilger Ltd, Bristol, 1985). ISBN 0-85274-427-7
B.W. Petley, Thirty years (or so) of the SI. Meas. Sci. Technol. **1**, 1261 (1990). https://doi.org/10.1088/0957-0233/1/11/023
T. J. Quinn, Metrology, its role in today's world BIPM Rapport BIPM-94/5 (1994)
A.P. Raposo, The Algebraic Structure of Quantity Calculus, (2016), arXiv:1611.01502v1
G. Rasch, On general laws and the meaning of measurement in psychology, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, (University of California Press, Berkeley, 1961), pp. 321–334
F.S. Roberts, Measurement theory with applications to decision-making, utility, and the social sciences, in *Encyclopedia of Mathematics and Its Applications*, vol. 7, (Cambridge University Press, Cambridge, 1985). ISBN 978-0-521-30227-2
G.B. Rossi, Measurement and probability – a probabilistic theory of measurement with applications, in *Springer Series in Measurement Science and Technology*, (2014), https://doi.org/10.1007/978-94-017-8825-0
T.D. Schneider, G.D. Stormo, L. Gold, A. Ehrenfeuch, The information content of binding sites on nucleotide sequences. J. Mol. Biol. **188**, 415–431 (1986). www.lecb.ncifcrf.gov/~toms/paper/schneider1986
M.M. Schnore, J.T. Partington, Immediate memory for visual patterns: symmetry and amount of information. Psychon. Sci. **8**, 421–422 (1967)
C.E. Shannon, W.W.Weaver The mathematical theory of communications. University of Illinois Press, Urbana, 117 p. (1963)
K.D. Sommer, B.R.L. Siebert, Systematic approach to the modelling of measurements for uncertainty evaluation. Metrologia **43**, S200–S210 (2006). https://doi.org/10.1088/0026.1394/43/4/S06
A.J. Stenner, M. Smith III, D.S. Burdick, Toward a theory of construct definition. J. Educ. Meas. **20** (4), 305–316 (1983)
A.J. Stenner, W.P. Fisher Jr., M.H. Stone, D.S. Burdick, Causal Rasch models. Front. Psychol. **4** (536), 1–14 (2013)
S.S. Stevens, On the theory of scales of measurement. Science **103**(2684), 677–680 (1946). New Series
L. Tsu, *Tao Te Ching, Chapter 71* (Vintage Books (Random House), New York, 1972)
J.A. Tukey, Chapter 8, Data analysis and behavioural science, in *The Collected Works of John A Tukey, Volume III, Philosophy and Principles of Data Analysis: 1949 – 1964*, ed. by L. V. Jones, (University North Carolina, Chapel Hill, 1986)
P.F. Velleman, L. Wilkinson, Nominal, ordinal, interval, and ratio typologies are misleading. Am. Stat. **47**, 65–72 (1993). https://www.cs.uic.edu/~wilkinson/Publications/stevens.pdf
T. Vosk, Measurement uncertainty: requirement for admission of forensic science evidence, in *Wiley Encyclopedia of Forensic Science*, ed. by A. Jamieson, A. A. Moenssens, (Wiley, Chichester, 2015)
J. Wallot, Dimensionen, Einheiten, Massysteme, in *Handbuch der Physik II, Kap. I*, (Springer, Berlin, 1926)
W. Weaver, C. Shannon, *The Mathematical Theory of Communication* (University of Illinois Press, Champaign, 1963). ISBN 0252725484
E.D. Weinberger, A theory of pragmatic information and its application to the quasi-species model of biological evolution. Biosystems **66**, 105–119 (2003). http://arxiv.org/abs/nlin.AO/0105030