CONTENT

CHAPTER I

The measurement problem from a conceptual perspective

Let me start these reflections on the measurement problem by drawing attention to the fact that measurement is a type of *description*. Description boils down to the translation of the 'raw' phenomena or 'objects' of investigation into *research data*, and this can be done in different ways.

First, there is *narrative* description, in which one describes the object of investigation with *words*. While at first sight such description seems to lack the level of exactness and sophistication required in science, it is nonetheless clear that it has played a major role in virtually every major scientific achievement, whether it be the work of Copernicus (1543), Linnaeus (1753), Darwin (1859) or any other groundbreaking scholar.

Second, there is description by means of (visual) images. Historical examples of this type of description include the detailed drawings of human anatomy by Vesalius (1543), the drawings and photographs of animals in the work of Tinbergen (e.g. 1948), and the x-ray crystallography pictures (Wilkins & Randall, 1953) on the basis of which Watson and Crick (1953) discovered the double helix of DNA.

Third, there is measurement or quantitative description. Measuring an object is essentially a matter of counting measurement units, or, in other words, situating this object alongside a unidimensional scale consisting of a measurement unit multiplied by a real number (Joint Committee for Guides in Metrology, 2012). This implies, strictly speaking, that only the unidimensional characteristics of objects can be measured. For instance, one cannot measure a car *an sich*. One first has to determine its unidimensional characteristics, such as length or weight, then choose a stable and invariant measurement unit, and finally count the number of units in the object characteristic to be measured. While common sense would seem to suggest that (sophisticated) measurement should offer a quasi-absolute assessment of an object, it is important to note that in reality it is always relative to the scale unit used. Mandelbrot (1967) illustrated this in a very tangible way in a paper entitled *'How long is the coast of Britain?'* If one looks at the British coastline, one does not see a straight line, nor even a smooth, curved line, but rather a twisting, meandering line. Measured with a scale unit of 200 km, the length of this line is 2400 kilometers; with a scale unit of 50 km, the length become 3400 kilometers (Figure 1). Mandelbrot was able to demonstrate that *"as the scale becomes smaller, the measured length of the coastline rises without limit, bays and peninsulas revealing ever-smaller sub-bays and sub-peninsulas"* (Gleick, 1987, p. 96). Richardson (1961) underlined the manner in which this can lead to practical problems. He checked encyclopedias in Spain, Portugal, the Netherlands and Belgium and found that the reported length of the common borders showed discrepancies of about 20%, due to different scales used.

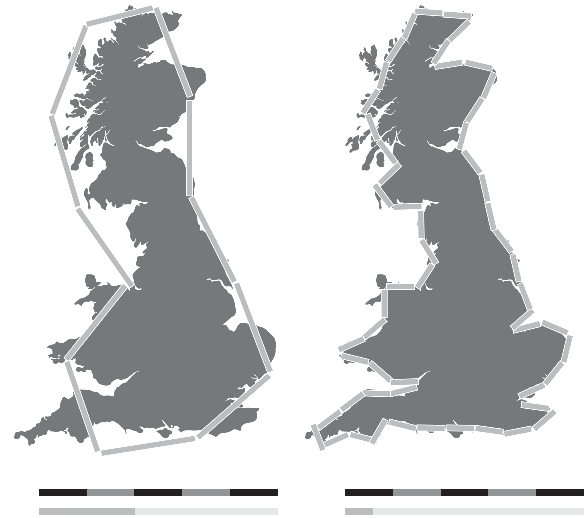

FIGURE 1
The measured length of the coast of Britain is 2400 km (approx.) with a scale unit of 200 km (left side of figure) and 3400 km (approx.) with a scale unit of 50 km (right side of figure) (image retrieved from Van de Sande, 2004).

Fourth, there is mathematical description (e.g. Prusinkiewicz & Fowler, 1995, p. 163), which is often overlooked or confused with quantitative description and which consists of determining the *formula* of a phenomenon. The Fibonacci formula, for instance, determines a spiral-shaped curve, which shows the proportions of the golden ratio and exactly matches a host of curves in nature, such as those found in sunflowers, seashells, tornados and galaxies. We could give numerous other examples; for instance, of fractal formulas that determine a variety of highly complex natural forms, such as seashells (Meinhardt, 1995), seaweeds, sponges and corals (Kaandorp & Kübler, 2001), and a variety of plants (Prusinkiewicz & Lindemayer, 1990). While the first three types of description are ultimately relative and imprecise – narrative description is possible using many different combinations of words, photographs or pictures can be made from many different perspectives, and measurement is always dependent on the scale used – mathematical description alone seems to touch the absolute. There is just a single mathematical formula to determine a mathematical shape. This gives mathematical description a special aura and status, making it reminiscent of Plato's universal forms, as if it is the finishing touch, the crowning glory of the scientific process.

Nevertheless, it is clear that all four types of description are necessary for effective science. Mathematical description would be a meaningless abstraction without the other descriptors. The ultimate objective of science must therefore be to 'knot' together the different types of description in constant reciprocal reference. The more solidly they are knotted together, the more consummate the scientific theory. In many respects, such a conception is preferable to a conception in terms of *explanation*. Explanation is always *causal* explanation. However, the ultimate cause of things regresses without limit. That is why science, in the final analysis, fails to explain. Something similar was voiced by von Neumann (1955): *"The sciences do not try to explain; they hardly even try to interpret. They mainly make models. By a model is meant a mathematical construct, which, with the addition of certain verbal interpretations, describes observed phenomena."*

Deciding which type of description is appropriate depends on the characteristics of the objects to be described. Consequently, we should ask ourselves: what are the objects of psychology and what characteristics do they possess? We can distinguish between numerous psychic phenomena, such as emotions, cognitions, memory, coping mechanisms, psychological disorders, personality traits, etc. That being said, there is a certain tendency in psychology to refrain from studying psychic phenomena and to

focus instead on external correlates at the biochemical, behavioral or neural level: *"[ever since] … the psyche arrived in the role of the object of investigation, everything has been done … to deny its reality and its status as the appropriate subject matter of psychology. Psychology is a science that has, quite successfully, denied its own existence. It has tried to get rid of itself by offering itself to physiology, sociology, computer science, and currently to neuroscience. Psychology is a science that is afraid of itself as a science!"* (Valsinere, 2012, p. xiii, cited in Schwarz, 2014, p. 214). It is clear that this has the potential to lead us astray, if only because, when we want to know whether a certain biochemical or neural process is indicative of a certain psychic phenomenon, we first and foremost need to define and describe the phenomenon itself. Therefore, in the broadest sense, the core objects of psychology are always what we might call the *psychic apparatus* and *psychic experience*.

Many perspectives can be taken on the psychic apparatus, yet for the purpose of this book I will refer to a *network* conception. Virtually all major psychological theories – from the early conceptions of Mesmer, Fechner, Janet, Charcot and Freud (see Ellenberger, 1970, pp. 145-149) to more recent cognitivist theorizations (e.g. Collins & Loftus, 1975) – maintain such a conception at one level or another. The core structure of the network is always an associated set of phonological-linguistic nodes (e.g. Stella, Beckage, & Brede, 2017) which, in turn, is associated with perceptual images (e.g. Bernstein, Clarke-Stewart, Roy, et al., 1994, p. 289) and bodily experiences (e.g. Galetzka, 2017; Bernstein et al., 1994, p. 308). While certain network aspects are the same across smaller or larger social groups – for instance, across everybody speaking the same language – others are based on a subject's historical and idiosyncratic set of experiences and are therefore subject-specific.

The network apparatus can be activated by triggers as varied as stimuli arising from memory, sensory images from the outer world, or the speech of other subjects. Whatever its origin, as soon as a trigger activates a node, the activation spreads in complex ways throughout the network and activates chains of phonological nodes, which manifest as internal thoughts that might or might not be externalized as speech. And as these chains of phonological entities manifest in subjective experience, so layers of *meaning* arise underneath. As early as 1890, one of the founding fathers of psychology, William James, remarked that meaning typically arises *"in pulses, after clauses or sentences are finished"* (James, 1890a, p. 182) and that the meaning of a phonological unit, of which the most prototypical example is a word, *"… consists of other words aroused, forming the so-called definition"* (James, 1890a, p. 160). For instance, the node 'hit' has a different meaning in the chain of words 'She *hit* him' in comparison with 'Her song became a *hit*' or '*Hit* the road'.

This all implies that the emergence of meaning is highly volatile and unpredictable. Minor differences at the level of incoming stimuli can provoke radical shifts at the level of meaning and psychic experience. Throughout this book, I will also refer to the distinction between the phonological nodes in networks and the meaning arising from the network by means of the more appropriate conceptual distinction between signifier and *signification*, as expounded in the linguistic theory of Ferdinand De Saussure (1906, p. 75). The term 'signifier' (i.e. a unit that *signifies something*) refers to the set of phonological entities or acoustic images that form the material basis of language (see De Saussure, 1906, p. 75), whereas the term 'signification' refers to the set of signifiers that defines another signifier.

As activation spreads throughout the network and meaning emerges, network nodes consisting of perceptual images (e.g. visual images, James, 1890a, pp. 360-361) and bodily experiences (e.g. sensations or vital impulses, James, 1890b, pp. 301-315) are also activated, so that rich, embodied, subjective experience occurs. The network actually delivers the total set of units a subject can use to attribute meaning (also see Bernstein et al., 1994, p. 307) and to generate psychic experiences. As such, it represents not only the basic structures of *memory*, but is also deemed to determine the set of psychic attributes that constitute what is usually called the *personality* of a subject, including emotional and affective responses, cognitive styles, motivation, social behavior, coping strategies and typical ways of appraising stimuli (Wells & Matthews, 1994), as well as more vital response tendencies such as aggression (Bushman, 1996), psychological symptoms, complaints and disorders (Bringmann, Vissers, Wichers, et al., 2013). Nearly all objects of psychological research can thus be situated within a network discourse, which justifies using it as a central reference point throughout this book.

This description of the psychic apparatus and the way in which psychic experience and meaning arise in it, shows us that the psychic system, in all respects, is a complex and dynamic system. This is nothing new; many others have come to the same conclusion (for example, Guastello, Koopmans, & Pincus, 2009; Salvatore, 2016, pp. 236-240). It is interesting to consider the consequences of this for descriptive methodology, starting from the dichotomy between mechanical systems on the one hand and complex dynamical systems on the other hand. A mechanical system interacts with its environment in a linear, invariant and predictable way according to a set of mechanical laws. The behavior of a ball that drops to the ground, for instance, is determined by mechanical laws that predict how it will accelerate, hit the ground, how high and in what direction it will bounce, and so on. In this respect, mechanical systems – from the simplest to the most complicated – are characterized by *deterministic predictability*.

In contrast, complex dynamical systems possess the most fascinating characteristic of deterministic *unpredictability*. A well-known example is convection rolls, which emerge when liquid or gas is heated. As the substance is heated, the rolls first behave in a stable, repetitive way. But when the temperature increases and reaches a critical limit, the rolls start to behave in a chaotic way. Interestingly, the behavior of the rolls in this phase can be described by a set of three non-linear equations with three unknowns and hence is strictly deterministic (Lorenz, 1963). At the same time, however, these mathematical equations never reach a final, stable solution and, consequently, the outcome is also non-periodic and fundamentally unpredictable. This unpredictability can also be understood from a more pragmatic perspective, based on the property of complex dynamical systems to display *sensitivity to initial conditions*, which means that the impact of the smallest factors can lead to radical differences in the way the entire system manifests itself. Since decisive differences can be infinitely small, measurement of the conditions can never be accurate enough to predict how the system will behave (Gleick, 1987, p. 19). A difference in the behavior of one liquid molecule at the beginning of the heating, for instance, can lead to the manifestation of radically different convection rolls. It is clear that the characteristics of deterministic unpredictability also apply to psychic experience and to the phenomenon of meaning. Psychic experience is strictly determined, in the sense that meaning does not arise in an arbitrary way. But because even the smallest details of a situation can lead to radical shifts in the meaning given to it by

a subject – the blink of an eye, a faint smile or an almost inaudible cough can radically change the meaning one attributes to another person's entire discourse – it can never be fully predicted what meaning will eventually emerge (Salvatore, 2016, p. 239).

The example of convection rolls is also illustrative of two other characteristics of complex dynamical systems that are particularly relevant for our line of reasoning. First, complex systems tend to *coalesce with their environment* (e.g. Bishop, 2011, p. 115), as if system and environment are both contained within an overarching supra-system. It is hard to determine, for instance, whether the behavior of a convection roll is due to the dynamics of the roll itself or to the currents in the liquid or gas in which it manifests. At the same time, the roll *organizes and is organized by* its environment. In a quasi-identical way, it is hard to determine whether a subject's behavior is due to the subject itself or to its social-psychological environment. In empirical research, this is illustrated by so-called researcher-demand or experimenter-expectancy effects, referring to the impact of the researcher's expectations and intentions on participants' responses. This has been described in papers with illustrative titles, such as *'Telling what they want to know'* (Norenzayan & Schwarz, 1999) and *'Self-reports: how the questions shape the answer'* (Schwarz, 1999). Danziger (1990, p. 8) rightly remarks that, although psychology now recognizes this problem and has even coined terms to describe it, nothing has really happened to change its methodology and develop designs in which subjects can be truly studied in their co-existence with other subjects. Moreover, it is remarkable that the so-called major sources of 'bias' in psychological measurement – such as acquiescence, compliance, social desirability and impression management – often refer to the profound impact of others on the responses of the participant. Consequently, this implies that psychological research systematically labels one of the fundamental characteristics of the object it intends to investigate as *bias* (!).

Another pertinent characteristic of complex systems is *historicity*. In sharp contrast to mechanical systems, the behavior of a complex system depends on its entire historical trajectory. Convection rolls, for instance, behave with a certain 'historical awareness', in the sense that the way in which the rolls structurally develop and manifest depends on all previous behaviors (Yin & Herfel, 2011, p. 407). This implies that descriptions of single aspects of complex systems inevitably have to refer to the non-local, historically developed, larger scale structures of the system (which also relate to the characteristic of hierarchy, i.e. the coherence between the parts and the whole of complex systems, Bishop, 2011, p. 114). However, 'the entirety of the system' is intrinsically multidimensional. It is composed of a variety of characteristics, which renders it unsuitable for metric description and necessitates narrative and/or mathematical description (Schmidt, 2011, p. 247). We have already argued from a network perspective that the psychic system – almost self-evidently – displays the characteristic of historicity. This can be observed in a very concrete way in the measurement context. For example, if you confront a participant with an item on a questionnaire that asks 'Do you feel depressed?', the response to this item will not only reflect 'the amount of depression' in this person, but also, to some extent, the participant's entire historically formed network that comes into play to make possible the interpretation of that item (Rosenzweig, 1933, p. 337). One subject might interpret the term 'depression' in a differentiated and sophisticated way, because the subject was raised by parents who worked in mental health care, while for another subject this same term might only evoke a global image of psychological distress. We will describe in the subsequent chapters how

this problem has been recognized by some scholars and how some of them have attempted to develop quantitative methodologies to solve it (e.g. Steyer, Mayer, Geiser, & Cole, 2015). In accordance with the line of reasoning presented here, we nevertheless believe that any thorough scientific inquiry implies that at some point the multidimensional, historical entirety of the subject's meaning-making apparatus is narratively described.

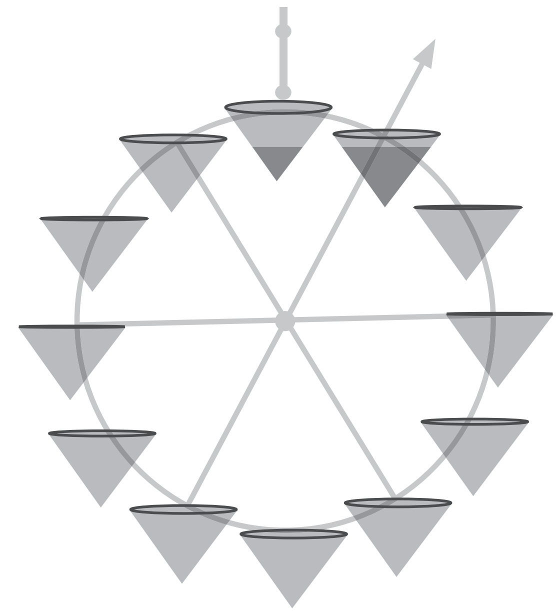

FIGURE 2
Lorenz's chaotic waterwheel (image retrieved from Leys, Ghys, & Alvarez, 2003).

The field of complex dynamical systems has yielded splendid examples of mathematical description. Perhaps the most illustrative is Edward Lorenz's famous chaotic waterwheel. A chaotic waterwheel is a wheel that has buckets with a hole in the bottom fixed to it, with water pouring in at the top (Figure 2). When the initial volume of water is small, the wheel moves at a constant pace in the same direction. However, as the volume increases and reaches a critical limit, the wheel starts to behave chaotically; its speed becomes irregular and the direction of rotation changes in an unpredictable way. Lorenz noticed that in many respects the waterwheel behaved like convection rolls and he was able to describe the chaotic movement at each moment with a set of three non-linear equations with three unknown variables x-y-z. Because the wheel behaves non-periodically, the values for x-y-z never stabilize and the same combination of values is never exactly repeated. In a moment of genius, Lorenz had the idea to plot the successive values of the three variables as x-y-z values in an orthogonal axis system. In an astonishing way, an esthetically magnificent (see also Ruelle, 1980, p. 137) butterfly picture appeared, which became known as the Lorenz attractor (Figure 3). In this way, Lorenz was able to *visualize* the order beneath a chaotic phenomenology by representing it in what later became known as *phase space*. Gleick (1987, p. 135) describes it like this: *"Phase-space portraits of physical systems exposed patterns of motion that were otherwise invisible, as an infra-red landscape photograph can reveal patterns and details that exist just beyond the reach of perception."* This is also reminiscent of what Husserl (1907, p. 141) called 'the pre-empirical', which is a space that pre-exists empirical space and determines it. And as previously remarked above, it also has much in common with Plato's universal forms that underlie the world as we experience it.

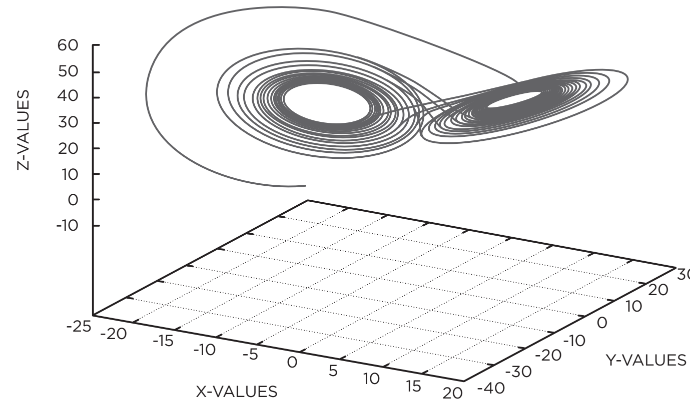

FIGURE 3
The Lorenz attractor (image retrieved from Reed, 2014).

This framework can now be used to evaluate contemporary measurement practice with reference to the nomothetic and the idiographic paradigms in psychology.

The differences between the two approaches were well summarized by Watson (1934) in his paper in *Psychological Bulletin*. The idiographic paradigm focuses on single case studies and considers measurements to be helpful, but only if they are situated in the context of the narrative descriptions of every single subject. Consider in this respect the following quotation from William Stern (1921, pp. 3-4, cited in Lamiel, 2009, p. 72), the inventor of the Intelligence Quotient: *"For the examinee in question, tests yield a number on the basis of which that examinee can be located somewhere along a quantitative scale, but which obscure things qualitatively peculiar to that individual. The results of direct observation cannot be quantified but make possible a qualitative refinement of the psychological profile. For all of these reasons, the methods of direct observation of an examinee must always be used to supplement the test methods, and the former must be developed with the same care*