Clinical Psychometrics
I attempted to effect the scientific in my psychopathology by methodological investigations, not by a dogmatic exposition of a complete psychiatric epistemology.

Karl Jaspers (1950)

The debt of psychiatry to the psychologist is now great and growing. From [Eysenck's] rigorous inquiries, sustained and resourcefully developed over years, psychiatry stands to gain an impetus and accuracy in some essential matters which will advance it and reinforce the free play of clinical skill and insight.

Aubrey Lewis (1952)

Emil Kraepelin is probably the most outstanding psychiatrist who ever lived.

Max Hamilton (1978)

To Ole Rafaelsen, a man larger than life, and to Erling Dein who showed me how to use Occam's razor in psychopathology
Contents

About the author, ix
Preface, x

Introduction, 1

1. Classical psychometrics, 3
   Emil Kraepelin: Symptom check list and pharmacopsychology, 6
   Charles Spearman: Factor analysis and intelligence tests, 10
   Harold Hotelling: Principal Component Analysis, 13
   Hans Eysenck: Factor analysis and personality questionnaires, 15
   Max Hamilton: Factor analysis and rating scales, 20
   Pierre Pichot: Symptom rating scales and clinical validity, 23

2. Modern psychiatry: DSM-IV/ICD-10, 27
   Focusing on reliability, 27
   Focusing on validity, 28
   Quantitative, dimensional diagnosis, 29

3. Modern dimensional psychometrics, 32
   Ronald A. Fisher: From Galton’s pioneer work to the sufficient statistic, 32
   Georg Rasch: From Guttman’s pioneer work to item response theory analysis (IRT), 34
   Sidney Siegel: Non-parametric statistics, 38
   Robert J. Mokken: Non-parametric analysis for item response theory (IRT), 39

4. Modern psychometrics: Item categories and sufficient statistics, 43
   Rensis Likert: Scale step measurements, 43
   John Overall: Brief, sufficient rating scales, 45
Contents

Clinical versus psychometric validity, 48
Item-response theory versus factor analysis, 49
Jacob Cohen: Effect size, 50

5. The clinical consequence of IRT analyses: The pharmacopsychometric triangle, 53
   Effect size and clinical significance, 53
   The pharmacopsychometric triangle, 56
   Antidementia medication, 59
   Antipsychotic medication, 60
   Antimanic medication, 65
   Antidepressive medication, 66
   Antianxiety medication, 69
   Mood stabilising medications, 72
   Combination of antidepressants, 73

6. The clinical consequence of IRT analyses: Health-related quality of life, 74
   The WHO-5 Questionnaire, 78

7. The clinical consequences of IRT analyses: The concept of stress, 82
   Post-traumatic stress disorder, 82
   The work-related stress condition, 84
   Integration of Selye’s medical stress model, 85

8. Questionnaires as ‘blood tests’, 89
   Population studies in depression and anxiety, 89
   The predictive validity of WHO-5, 92
   Screening scales, 92

9. Summary and perspectives, 95

10. Epilogue: Who’s carrying Einstein’s baton?, 103
   Glossary, 109
   Appendices, 114
   References, 185
   Index, 196
About the author

Per Bech

Per Bech received a medical degree from Copenhagen University in 1969. In 1972 he received a gold medal award from Århus University for his thesis on the dose-response relationship between cannabis (tetrahydrocannabinole) and various psychological measurements, including time experience and reaction time in simulated car driving.

He completed a doctorate thesis (Dr. Med. Sci) at Copenhagen University on the clinical and psychometrical validity of rating scales in depression and mania in 1981.

He was appointed Professor of Psychiatry at Odense University in 1992 and in 2008 he was appointed Professor of Applied or Clinical Psychometrics at Copenhagen University.

Since 1981 he has held the post of chief psychiatrist at The Mental Health Centre North Zealand in Hillerød (Capital Region of Denmark) and is Head of the Psychiatric Research Unit there. He is an honorary member of the Royal College of Psychiatrists and of the European Psychiatric Association (EPA).
Preface

The first edition of this book was the original Danish version published in January 2011, as an introduction to the very broad field covering clinical psychology, psychiatry and clinical psychopharmacology. It was an attempt to follow Kraepelin’s rating scale approach and his pharmaco- psychometrics as they have developed in the twentieth century, especially with the introduction of psychopharmacology in the 1960s. The central concept here is the Pharmacopsychometric Triangle, in which (A) covers desired clinical effect, (B) unwanted effects, or side effects, and (C) patient-reported quality of life. In connection with (A), short psychometric scales are described which can be used to measure such classes of drugs as antidementias, antipsychotics, antimanics, antidepressants, antianxiety drugs, and mood stabilisers.

The psychometric performances of scales for (A), (B) and (C) are described with reference to both factor analysis and to item response theory models. These models have been amended for readers without mathematical knowledge. However, throughout the book experienced psychiatrists are referred to as an index of validity in an attempt to bring the symptoms home to the dimensions within (A), (B) and (C) where they belong.

My thanks when preparing the Danish version of my book went, as so often before, to Peter Allerup, Professor of Theoretical Psychometrics at the University of Århus. He has been a ‘basic factor’ for my work with rating scales over nearly 40 years! My research coordinator Lone Lindberg has made a unique contribution, with invaluable help in typing and layout. Gabriele Bech-Andersen and Susan Søndergaard are behind the translation procedures for the scales in the Danish version, and Susan has translated this English version from the Danish. Ove Aaskoven has been my statistical research assistant for many years, often in collaboration with Peter Allerup. Finally, I owe a debt of thanks to the Munksgaard editors Marie Schack and Daniel R. Andersen who made helpful suggestions for the earlier Danish versions.

In this English version editor Jesper Konradsen has raised challenging queries, especially on the philosophical lines running through it, with
focus on the development of psychometrics from a philosophical start to mathematical aspects of measuring mental stages, to clinical validity and dose–response relationships and then back to the philosophy of Wittgenstein, which brings symptoms home to form relevant syndromes or dimensions.
Clinical psychometrics has incorporated psychology as an important auxiliary subject in the same way as neuropharmacology and neuroanatomy. As a branch of medicine, clinical psychiatry has especially attempted to determine the organic cause of mental disorders; and before the establishment of psychometrics, the psychological approach to patients was seen as a non-organic explanatory model for mental disorders. Freud's psychoanalysis, in particular, was seen as a psychological explanatory model; partly because psychiatry was regarded for many years as an atypical branch of medicine due to the non-testability of the Freudian theories, which were thus without clinical validity (1).

The scientific approach to psychology launched by psychometrics has resulted in psychiatry being regarded as a clinical branch of medicine. This only took place with the 1987 publication of Feinstein's monograph on clinical psychometrics (2). Finding a comprehensive overview of the role of psychometrics in clinical psychiatry has proved difficult. The following is an attempt to put this to rights.

It falls naturally to divide clinical psychometrics into two eras. The first of these, the classical era, covers the period from 1879 to 1945. It is the era of the greatest names: Wilhelm Wundt who founded psychometrics in 1879 and his two most important pupils; Kraepelin and Spearman. The modern period developed after 1945 has Eysenck, Hamilton and Pichot as the major psychometricians. They developed the questionnaires and rating scales archetypal of modern clinical psychometrics in the period from 1945 to the 1970s (3). From a statistical point of view, however, Francis Galton and his London psychometric laboratory (founded in 1884) are essential elements, together with Galton's two most important 'students' (Pearson and Fisher) and the three people (Rasch, Siegel and Mokken) who developed the psychometric analyses that are
archetypal of modern clinical psychometrics in the period from 1945 to the 1970s (4) (see Figure I.1).

The most obvious impact of modern psychometric research, which has resulted in short valid rating scales and the descriptive statistics of effect sizes, is the pharmacopsychometric triangle. It was the revolution in pharmacology 50 years ago that led to the rebirth of Kraepelin’s pharmacopsychology, now crystallised in the pharmacopsychometric triangle, the major focus of this book.
More than a century ago, psychology was defined as the science of human mental manifestations and phenomena. However, it was psychometrics (the science of measuring these mental manifestations and phenomena) that made psychology scientific. Thus, psychometrics is a purely psychological area of research.

From a historical point of view, psychology branched out from philosophy as an independent university discipline at the close of the nineteenth century. It all started in Leipzig in 1879. Here the philosopher Wilhelm Wundt (1832–1920) established his psychological laboratory at the university. Formally, however, his laboratory remained under the faculty of philosophy. Wundt succeeded in detaching psychology from philosophy, especially freeing it from the influence of Emanuel Kant, an extremely influential philosopher who stated that it is impossible to measure manifestations of the mind in the same way as physical objects (5). With his criticism of pure reason, Kant (1724–1804) established the very important distinction between ‘the essential nature of things’ (things in themselves) and ‘things as they seem’ (i.e., that which we sense or perceive as a phenomenon when faced with the object we are examining).

Figure 1.1 illustrates Kant’s philosophical approaches with reference to present day psychiatry, according to which depression is understood to be a clinical phenomenological perception (shared phenomenology of depressive symptoms) as measured by the six depression symptoms contained in the Hamilton Depression Scale (HAM-D<sub>6</sub>, see Figure 3.1). Modern neuropsychiatry attempts to describe the depression behind the phenomenological perception, i.e., depression ‘in itself’, as we believe it to be present in the brain, for example, as a serotonin 1A receptor problem (impairment).

The area of research now known as brain research is just such an attempt to measure the processes presumed to be taking place in the brain, that is ‘das
Ding an sich’. As pointed out by Sontag, reality has increasingly grown to resemble what the camera shows us (6). It is reality itself when the neuropsychiatric camera demonstrates receptor binding in the brain, while clinical reality is increasingly becoming what the camera visualises for us by means of assessment scales or patient-related questionnaires.

The ability to describe reality as it is in itself, i.e., looking at the world unclouded by any preconception of it, has been debated by such neo-Kantentians as Wittgenstein and Quine (7). The quantification of endophenotypes or deep phenotypes is probably the most scientific image of the world. However, we do not have endophenotypes to tell us whether we indeed can describe reality, e.g., the brain, as it is itself. Wittgenstein tells us that he does not want to say whether we can or cannot describe reality as it is in itself. He wanted, as stated by Putman to bring our phenomenological items back to their home in clinical psychiatry. This is what clinical psychometrics is about (7).

Figure 1.2 shows a correlation between the so-called psychotic symptom items in an American rating scale (see Appendix) and serotonin 2A receptor binding, which it is now possible to measure by means of positron emission tomography (PET) scanning (8). The figure shows a correlation coefficient of $-0.57$; this is statistically significant but not clinically significant, as the variance on the ordinate axis (the ‘psychosis’ scale) can explain only about 32% of the variance on the axis of abscissas (serotonin 2A receptor binding). If the two patients at the far left are excluded as outliers, then the negative correlation value is halved, so that less than 10% of the variance is explained.
The scale in Figure 1.2 shows the positive symptoms in a schizophrenia scale. In the early 1970s, the American psychiatrist Nancy Andreasen found it important to label those schizophrenic symptoms on which medication had an effect as positive. In clinical psychiatry, these were termed productive symptoms as they were often the reason for hospitalisation in a mental institution. Later on, Nancy Andreasen became interested in neuropsychiatric brain imaging methods [Computer Assisted Tomography (CAT scan), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET)], which became available in the 1980s and 90s. However, in an interview from 2003, she had to admit that schizophrenia is probably not located in one specific section of the brain (9). Schizophrenia affects many different brain areas that cannot be visualised as ‘das Ding an sich’.

Wilhelm Wundt’s major achievement was to realise that mathematical models of ‘das Ding für uns’ can be used to measure the ‘shared phenomenology’ of the state one wishes to assess quantitatively. During his studies at the Heidelberg faculty of medicine, he obtained a degree in medicine. Wundt then participated in studies in the physiology of perception under Helmholtz (1821–94) and Fechner (1801–87). He observed that it was possible to get subjects to reliably assess sensory impressions when the conditions of the study were standardised, e.g., with increasing light or noise exposure.

**Figure 1.2** The problematic relationship between the clinical, the psychometrical and the biological frames of reference with a correlation coefficient of −0.57
Wundt’s philosophical basis was that each manifestation of the mind corresponds to a neurobiological substrate in the brain, but in his opinion the psychometric measurement of this manifestation of the mind should only focus on the psychological phenomena (das Ding für sich) and not include any biological elements in any way. He belonged to the branch of philosophy called non-reductive monism (corresponding to Harald Høffding’s critical monism, which maintains that manifestations of the mind cannot be reduced to purely biological variables) (10). On the other hand, it is of course possible to reduce certain manifestations of the mind to less complicated ones in an attempt to obtain the most reliable or objective measure. He felt that it would be possible in this way to make psychology scientific within the frame of its own descriptive realm, since psychological and biological methods of description are two different ways of viewing reality.

Wundt’s approach was that of descriptive psychology where the various dimensions consisting of individual items (symptoms) can be added to give a total score. He was excluding the immediate, peak-experiences detached from relations, e.g., the spontaneous, stimulus-unrelated, perception-like images in the religious experience of the child, actually referred to as ‘Sensus numinis’ (11,12). The clearest description of Wundt’s scientific approach based on his ‘Grundzüge der psychologischen Psychologie’ is found in Vannerus’ monograph (13).

The psychometric method developed by Wundt is probably the only specific psychological method identified in mental science, i.e., in scientific psychology (14). The two most famous scientists to emerge from Wundt’s psychological laboratory in Leipzig were Emil Kraepelin and Charles Spearman; both of them understood that psychological measurement (psychometrics) and biological measurement are two different ways of viewing nature.

**Emil Kraepelin: Symptom check list and pharmacopsychology**

Kraepelin (1856–1926) had just obtained his medical degree when he applied for a post at Wundt’s laboratory in 1882. As Wundt was unable to finance his salary, Kraepelin also had to take up a post as a locum at the local mental hospital in Leipzig. Thus, Kraepelin held an unsalaried position at the Wundt laboratory. Kraepelin’s purpose was to introduce scientific psychology into psychiatry so that his career as a psychiatrist would be furthered by his studies at Wundt’s psychological laboratory. In his job application to Wundt, he wrote that he would give a kingdom for a [research] topic; Wundt then gave him the opportunity to examine the influence of psychoactive substances
Kraepelin’s symptom checklist from his Zählkarten (counting cards)

- Nervousness
- Restlessness
- Irritability
- Depression
- Psychomotor retardation
- Aggression
- Grandiosity
- Negativistic behaviour
- Hallucinations
- Paranoid ideas


Figure 1.3 The assessment scale or checklist used by Kraepelin (10)

such as alcohol and the hypnotic drug chloral hydrate on volunteer research subjects. Kraepelin set out to demonstrate a dose response curve using reaction time measurements as the psychological response and psychoactive substances as the stimuli, so that increasing amounts of alcohol (number of drinks) led to lengthening reaction times. Since Wundt could see that Kraepelin had his heart set on psychiatry, he encouraged Kraepelin to employ this objective scientific method when subsequently assessing the various symptoms presented by patients suffering from mental disorders.

Kraepelin published his first Psychiatric Compendium as early as 1883. In this he attempted to focus on the symptoms presented in the different disorders (Compendium der Psychiatrie. Verlag von Amber Abel, Leipzig, 1883). After leaving the Leipzig laboratory and starting on his career as a psychiatrist in Munich, Kraepelin published several compendiums or textbooks on psychiatry. He revised his textbook almost bi-annually and in the 6th edition in 1899, he was able to describe two disorders with different symptom profiles: manic-depressive disorder and schizophrenia.

Figure 1.3 shows the checklist Kraepelin used when systematically monitoring his patients over several years in order to ascertain which symptoms possessed ‘shared phenomenology’ over this period of time. These are called
checklist symptoms, as Kraepelin only determined whether the symptom was present or absent. This type of scale is called a nominal scale. Using this method, Kraepelin was able to demonstrate that during a period of about six months, some patients presented with the first five or six symptoms in Figure 1.3, while in other episodes of shorter duration (up to three months) they had the next two symptoms (aggression and delusions of grandeur), along with restlessness, sleep disturbance and irritability. Between these episodes of depression or mania, these patients were discharged from hospital and were socially well-functioning. Other patients, who were often lifetime residents in asylums, had the last three symptoms in Figure 1.3. Kraepelin described them as suffering from dementia praecox (now schizophrenia), as the disorder typically started when they were about 20 years of age and was chronic in nature, often with an influence on intellectual functions as well. But these were consequences, not elements, of the schizophrenic symptomatology. Manic-depressive disorder, on the other hand, did not typically emerge at a specific age. Based on the original registrations by Kraepelin on his ‘Zählkarten’ (counting cards) including the checklist symptoms in Figure 1.3, Jablensky et al made a comparison using the Present State Examination (PSE). From the PSE scores the ICD-9 diagnoses of schizophrenia and manic-depressive disorder can be made. In total Jablensky et al identified 721 patients assessed by Kraepelin and found a concordance for the diagnoses of schizophrenia and manic-depressive disorder of approximately 80% with the ICD-9 diagnoses (15).

In his thesis: ‘Über die Beeinflussung einfacher psychischer Vorgänge durch einige Arzneimittel’ (Jena, Fischer Verlag 1892), Kraepelin established the area of research he designated pharmacopsychology.

In the 8th edition of his textbook, written between, 1909–13, Kraepelin added reflections on the psychotherapeutic effects of certain drugs such as morphine, phenemal and chloral hydrate. However, he found that the effects of these drugs on schizophrenia and manic-depressive disorder were extremely poor. He was thus able to observe the spontaneous course of illness in these two disorders.

In the schizophrenic patient, as stated previously, the condition was unremitting, while manic-depressive disorder was characterised by episodes with specific symptoms and then periods between episodes of a year or more in which the patients were completely without symptoms and thus able to function normally. In these descriptions, Kraepelin determinedly avoided including the various theories on disease circulating at that time, such as hereditary elements, stress burden and so on.

Kraepelin’s textbooks were not widely known outside Germany, as the two world wars made German psychiatry less acceptable. His system only began to make an international impact after World War II, not least in the USA.