Validating Concepts of Mental Disorder: Precedents from the History of Science.

Robert Miller

Lecture given: Victoria University, Wellington, Thursday 30th August 2012[1]

Psychiatry at present is attempting to define itself as a scientific discipline, with credentials equivalent to those in other areas of medicine. Much as I support this ambition, I believe that it is by no means fulfilled. In part this is because the task is more fundamental than in other areas of medicine; but, in direct consequence of this, psychiatry, I think, must go through stages which were crucial in the emergence of the natural sciences in the seventeenth century. There is little awareness of these stages in psychiatry. So, this talk will emphasise the history of science in areas far removed from psychiatry. Nevertheless, the talk is also highly relevant to the here-and-now, in that one of the central concerns of lay people, lacking the expertise of researchers in psychiatry, also points to a need for research a much more fundamental level than that on which most of it currently focuses. The subject matter then, in which diverse communities and deepest theoretical researchers have common cause, is the scientific status of our concepts of mental disorder.

Community perspectives

Psychiatrists freely uses the term "mental illness" but disease concepts in psychiatry seem to be rather fuzzy, based on conventions, sustained by prestige of people in authority and the faith of their followers. I have heard endless, fruitless debates about classification since the early 1970s, debates going back to the nineteenth century. The circularity of definition is even referred to in a line from Shakespeare's *Hamlet:* "To define true madness, What is't but to be nothing else but mad?".

Lay communities are well aware of these shortcomings. Their members are not experts in psychiatry, but they are experts in their own life experiences. What are their concerns? Here are a few examples:

[1] Omitting all overheads except one, many of which were portrait pictures of scientists mentioned in the text.

(i) It is the experience of many patients that they receive a variety of different diagnoses from different psychiatrists for one disorder. Ever-more emphatic claims by psychiatrists that "mine is the *correct* diagnosis", cuts no ice. It brings psychiatry into disrepute.

(ii) People in the community rightly ask: "Isn't it absurd that people be placed into mutually-exclusive, non-overlapping diagnostic boxes"? Surely human diversity requires something more subtle. There is concern that psychiatry is medicalizing human diversity rather than welcoming and celebrating it. There are real issues here about what constitutes mental disorder.

(iii) It is also suggested that psychiatric diagnoses serve commercial interests (e.g. health insurance and pharmaceutical industries), rather than needs of patients. Diagnoses seem to be "made up" to serve such interests, with no secure rational basis.

(iv) We have seen a major movement across a number of countries, to abolish the word "schizophrenia" as a diagnostic term. This is propelled in part by community concern that the diagnosis is stigmatising, that it is "a life sentence more than a diagnosis". This may split North American from British and European psychiatry.

(v) In some parts of the Western world, there is growth of the claim that "schizophrenia is not a disease", resurgence of anti-psychiatry rhetoric popular in the 1960s and 1970s, and rejection of gains from biological approaches. In some places this has undermined major aspects of mental health care (including therapy with antipsychotic medications). It alarms psychiatrists, as it alarms me, but the profession cannot always mount an effective defence. The "biological revolution" in psychiatry has not gained "grass-roots" support.

(vi) For another diagnostic entity - attention deficit/hyperactivity disorder (ADHD) - it is asked: Is it really a mental disorder? . . or is it a normal personality variant, which is a disorder only in certain social environments (especially those created in schools). Perhaps more attention should be given to unhealthy school environments as a public health initiative rather than treating ADHD as one for personal health care (and medication with ritalin).

(vii) Another diagnostic category - dyslexia - is certainly disabling, given that our culture relies heavily on the written word; yet it is well understood that people with dyslexia often have unusual talents in other areas, which enable them not only to hold their own, but even to achieve pre-eminence[2].

---

2 Joanne Black "*In their right mind*" New Zealand Listener, May, 8-14, 2010.

(viii) In Britain the government tried to foist the term "dangerous severe personality disorder" as a diagnosis, with neither a legal nor a medical basis, this to be used as a basis for pre-emptive detention of people who had committed no crime. The same was attempted in New Zealand, and was stopped only when key psychiatrists put their own jobs on the line over the issue. Such political interference with psychiatry is made easier because few of its *other* diagnoses have secure scientific status. It is good that there are people with sufficient integrity to stop it, but one cannot rely on that. One needs other safeguards.

(ix) In New Zealand, the government-backed campaign "Like Minds Like Mine", which I work with, aims to combat stigma and discrimination related to mental illness and has received acclaim around the world. Persons with lived experience of mental illness played a major part in shaping this campaign and its implementation, yet it *avoids* diagnostic labels, preferring instead to use direct first-person accounts of lived experiences. Thus, in some areas, the idea that diagnosis is essential to define mental disorders and to guide their treatment is being overtaken by events, and by public awareness.

All this points to real problems about the status of many concepts of mental disorder used in psychiatry. What has gone wrong? Is there a basic misconception? If so, what is it? To address these issues, I want to start from two basic dichotomies. The first goes back to medieval times, the split between two approaches to the study of the natural world (precursors of science). The other is the distinction between experiment and theory as methods of exploring the natural world.

Natural Philosophy versus Natural History.

Before the birth of the natural sciences, their precursors were two areas of scholarship, *natural philosophy* and *natural history*. The aim of *natural philosophy* was to *explain* natural phenomena, using reasoning from either natural or supernatural assumptions. This developed into what we now call physics, and the approach spread to other areas (chemistry, biophysics, etc). The role of *natural history* was different: to *describe* nature as it appears in all its complexity. In origin it was qualitative, but later could be quantitative. Correlation and association are part of this, being aspects of description, not to be confused with explanation. The key difference between the two is that natural philosophy - that is physics - deliberately simplifies what it studies, so that very few variables are relevant. One then has a chance to explain things. Natural history deals with the natural world, life, and history in their full complexity, with a wealth of descriptive detail, which is then far too complex to find fundamental principles for explanation or cause.

Experiment versus Theory.

The second dichotomy applies mainly in the natural philosophy tradition, the distinction between *experiment* and *theory* (or, if you like, between ideas, linked by reasoning, and empirical observations). Before the seventeenth century, for two thousand years in the Western world, two approaches, empiricist and rationalist, had long been in rivalry, as ways to discover truth. Mainly the rationalist approach was dominant, because of the power of the Catholic church. In the seventeenth century, for the first time, the two started to combine. Empirical observations were sometimes descriptive, but later came increasingly from systematic experiments. Reasoning, from the time of Galileo, tended to be quantitative and mathematical, although that is not a necessary part of the tradition.

The first era when this combination came about involved three pioneers from the sixteenth and early seventeenth century. Profiles of these three give insight into the relation between theory and experiment.

Copernicus

Nicholas Copernicus (1473-1543) a polymath from Northern Europe did little *observational* astronomy, but is famous for the proposal (published on his deathbed) of the heliocentric view of the solar system. This was neater mathematically than the Ptolomaic system inherited from classical periods, but needed little new empirical data.

Tycho Brahe

Tycho Brahe (1546-1601) was a man of totally different stamp. He was a Danish Nobleman, and while a student in Leipzig in 1563, saw the alignment of Saturn and Jupiter one month away from the date predicted on the Ptolomaic system. The discrepancy led him to undertake systematic study of how planets *actually* moved. In due course he sought the help of the King of Denmark, who gave him an island in the Baltic, which he called Uraniborg. A palace was built for his work. He recruited assistants to make observations, all at night (this being before telescopes were invented), and others for calculations. He produced accurate data on planetary motion night-by-night, on clear nights, over a period of nearly thirty years, the best empirical data ever produced up to that point on any subject.

Kepler

One of his calculators was *Johannes Kepler*. From a humble background, in what is now south Germany, he studied at the university of Tübingen. He was imaginative, had a flair for maths, and secretly studied - and became convinced of - the Copernican system (although the university still taught the Ptolomaic system). When he met Tycho Brahe, Brahe respected Kepler's skill in computation, if not his belief in the Copernican system, and, in 1597, Kepler joined him at Uraniborg. They worked together for a few fractious years, before Tycho Brahe died, leaving Kepler with an abundance of high quality empirical data, upon which he could employ his imagination and mathematical skill. On his death-bed, Tycho is said to have pleaded with Kepler *not* to adopt the Copernican system. Kepler wouldn't have a bar of it, and soon made the momentous discovery that planetary orbits were not circular. This defied two thousand years of teaching since Aristotle. Further work revealed the mathematical system which *did* describe planetary orbits, first for Mars, then for other planets: They were elliptical, with the sun at one pole. Further study revealed the finding that the area swept out between an orbit and the sun was equal in equal times, despite change in velocity.

Diagram

He published this in 1609, and full astronomical tables following in 1627.

I offer some comments on these three pioneers: Copernicus and Kepler were both theoreticians, Brahe was an empirical scientist. From what one can gather the first two had totally different temperaments from Brahe - vivid in imagination, yet concerned with rigour in reasoning, whereas Brahe, less imaginative, had a dogged, perhaps obsessive concern to get the best possible data, blind to theory or explanation. The theoreticians needed little finance, and worked in isolation; Brahe needed big money and a big team. Kepler and Brahe, needed each other, but did not get on well together. Their different temperaments, and habits of thought were unlikely to be combined in one person. Relations between the two types are likely to be tense. Nevertheless this may be the first time that the rationalist and the empiricist approaches worked together in synergy; and the combination provided Isaac Newton with a start for his own monumental work, seventy years later.

Bacon

About this time we have the prophetic writings of *Francis Bacon* (1561-1626), a contemporary of Shakespeare, the first to write on the basic method of what we now call "science". He is often seen as the first modern advocate of empiricism. Given the two-thousand year dominance of rationalism since Pythagoras and Catholic philosophers, that is fair. But he actually advocated a measured balance between empiricism and rationalism. Here is a wonderful example of his elegant writing.

> "Those who have handled sciences have been either men of experiment or men of dogmas. The men of experiment are like the ant, they only collect and use; the reasoners resemble spiders, who make cobwebs out of their own substance. But the bee takes a middle course: it gathers its material from the flowers of the garden and of the field, but transforms and digests it by a power of its own. Not unlike this is the true business of philosophy: for it neither relies solely or chiefly on the powers of the mind, nor does it take the matter which it gathers from natural history and mechanical experiments and lay it up in the memory whole as it finds it, but lays it up in the understanding altered and digested. Therefore from a closer and purer league between these two faculties, the experimental and the rational (such as has never been made), much may be hoped. (from *Novum Organum,* 1620).

Since the time of these pioneers, interplay between ideas and experiments (between theory and observation) has been the cornerstone of research in the natural philosophy tradition. A form of reasoning emerged - of which there are many examples - which I would like to call "cross-level explanation".

Cross-level explanations

In this, arguments are presented by which phenomena known at a "higher level" are accounted for by simple premises about lower level processes. Often the premises are quite hypothetical, because they are beyond techniques currently available. Examples are the reasoning leading John Dalton to his atomic hypothesis in 1800, and later in the century, the formulation of the kinetic theory of gases, by which the gas laws were accounted for by the motion and collision of hypothetical things called molecules. Interaction between ideas and experiments is the *defining* characteristic of the natural philosophy tradition.

Interaction between ideas and experiments defines the natural philosophy tradition

----

In physics, theoreticians and experimentalists have tended to be different people, with different skills and attitudes, dependent on and respecting the skills and attitudes of the other - a synergy which has made progress in physics so rapid, and secure. One sees this at its best in the twentieth century, in the collaboration between Ernest Rutherford and Neils Bohr. A key point about such explanations is that, while one has to start from key empirical facts, one need not know *all* relevant facts. Indeed, predictions made about unknown areas allow critical tests of preliminary hypotheses.

What about the validation of the central concepts? Of course, precise reasoning requires precise concepts. Or, as Francis Bacon put it:

"If the notions themselves (which is the root of the matter) are confused and over-hastily abstracted from the facts, there can be no firmness in the superstructure."

In the natural philosophy tradition, explanation is closely linked to the validation of concepts. In natural philosophy, there were originally four key concepts: length, time, mass and force.

M,T,L.F.

Length *can* be defined precisely since it is easily and reliably measured. Introduction of time as a quantitative variable came more slowly, being absent in ancient Greek science. It was Galileo who first used "time" as a quantitative variable to explain empirical data. The critical step was however made by *Isaac Newton*.

Newton

Before Newton, the words mass and force had no proper definition, like the concept of schizophrenia today. It was the solid reasoning of Newton, involving the quantitative relations between length, time, mass and force which validated the definition of those concepts. His staggering achievement was to define terms in particular ways, and to devise a system of reasoning (mathematical reasoning, but it need not be mathematical) such that his scheme would *explain* many phenomena in the natural world. In more detail, "mass" was defined independent of weight, as "resistance to acceleration"; and "force" was then what causes acceleration (or deceleration), but not needed for uniform motion. The laws of motion and of gravity used these definitions, and explained planetary motion and many other things, with a precision never seen before. As a result the terms mass and force became concepts which *were* validated, in a strong way. Thus the basic language of the natural sciences was established, a language which, since Newton's days, has been extended, modified, and (in relativity theory) greatly deepened, but not fundamentally overturned. The language is valid in all countries and cultures, and crosses generations. So, "science" has world-wide appeal. Concepts like mass and force *do* have more precise definition than ones used in humanistic debate (such as "democracy" or "freedom"), and so reasoning is more precise.

Messages

There are several messages here: Explanation and validation of concepts are mutually interdependent. The *only* way in which scientific concepts can be securely validated is when they are defined in such a way as to support strong explanatory arguments. It is *exceedingly* difficult, because explanation depends on the way concepts are defined, but one doesn't know how to define the terms until the explanation is in mind. There is no short cut, no easy algorithm, no linear chain of reasoning bound to succeed; and, at risk of sounding like Margaret Thatcher, I assert there is no alternative. The process is circular: The conclusion depends on the premises and the premises depend on the conclusion. Difficult it may be; but when it works, it works like wildfire, and "feeds on itself".

Compare that with approaches to classification in psychiatry. The key figure is Emil Kraepelin, at the height of his powers in Germany 100 years ago.

Kraepelin

It is to him that we owe current concepts such as schizophrenia, mani-depressive illness and several other entities. He had four principles for classification:

1. Mental disorders are best understood by analogy with physical disorders.
2. Medicine's historic first step was to classify. Psychiatry must begin there also.
3. Classification of mental disorders demands careful observation of visible phenomena.
4. *Classification is a necessary first step to understanding aetiologies.*

I can raise questions about all of these. On the first point, I simply ask "Why?" On the second and third I point out the contrast with physics, where there is a much greater role for inference from raw data, and hidden variables. But on the fourth I draw the sharpest contrast with physics. For Kraepelin, classification must precede elucidation or understanding; in physics the two are interdependent. So, about Kraepelin's maxim, one has to ask: On what principles, and on what authority is classification to be conducted. Possibly on the hunch of an acknowledged expert or authority – perhaps Kraepelin himself. So it seems, looking back over 100 years. So, there you have it, in a nutshell: Scientific Reasoning versus Medical Authority. You pays your money and you takes your choice!!

Origins of biological science: Natural Philosophy versus Natural History:

Biological systems are more complex than those in physics, and not easily simplified to reveal single variables at work. So, biology has tended to be contained mainly within the natural history tradition, dealing with nature in its full complexity. The origins of biology, are essentially descriptive not explanatory. Medicine also originates in natural history; and in psychiatry, pioneers such as Pinel and Esquirol aimed to describe, not to explain. There have however been notable successes in biomedicine where something akin to natural philosophy was possible, including cross-level explanations; and often these breakthroughs have come from scientists with a background in the physical sciences. The germ theory of infectious disease (if not the discovery of specific infectious agents) may be the first such success. Modern examples include the unravelling of the ionic fluxes underlying the action potential in the 1950s, and, in the same era, the biggest of all such insights, the revelation of how the molecular structure of DNA could explain facts of reproduction of cells and organisms, and many facts from genetics. It is notable that two key figures in the latter revelation were Maurice Wilkins, with a physics degree from Cambridge University, and Francis Crick, who studied at the Cavendish physics laboratory there. Another pioneer of molecular biology, Jacques Monod, was clear he was working in the natural philosophy tradition, when he gave his book “Chance and necessity” the subtitle: “an essay on the natural philosophy of modern biology”.

Mainly however, in biology, and more so in medicine, the systems studied are so complex that description, not explanation has been the primary aim. Isolating the impact of single variables is difficult and often assumed to be impossible. If “explanation” is claimed it is of weaker kind than in the natural philosophy tradition. “Biological variation” is accepted without question, and submitted to statistical analysis. Rarely is it the object of explanation as it might be in physics, where most systems are exactly-reproducible. Thus statistics is important, and theoretical reasoning (whether quantitative or not) is rare. More typical in bio-medicine is the style of research formulated in the nineteenth century by the physiologist, Claude Bernard.

Claude Bernard

He aimed to establish the use of the scientific method in medicine. However, his concept of “scientific method” was not that used in physics. He writes

----

“Proof that a given condition always precedes or accompanies a phenomenon does not warrant concluding with certainty that a given condition is the immediate cause of that phenomenon. It must still be established that when this condition is removed, the phenomenon will no longer appear”. (from *An Introduction to the Study of Experimental Medicine,* Claude Bernard, 1865; English translation, published by Dover, 1957, p.55).

His criterion is an empirical one based on physiological experiments, not one based on exact reasoning, as in physics; the word “cause” is used differently from how it might be used in physics; and “proof” was less certain than in physics. One can make the same point about Robert Koch’s criteria for supposing a microbe to be the cause of an infectious disease, or Henry Dale’s criteria for showing a chemical substance to be a neurotransmitter. I am not decrying the progress made with this approach; but we should be aware of its limitations, and we can do better. The question I *do* want to ask is whether true explanation as in natural philosophy can be achieved in systems of the complexity of the human brain, and which underlie at least some aspects of disorders dealt with in psychiatry. I think we are not sufficiently ambitious: Real explanations *can* be discovered, based on evidence we already have. To substantiate this claim, I need to say more on the differences between physical and biological systems.

True explanation: Differences between biology and physics

For any scientific explanation to be successful, critical empirical facts must be known, and no crucial confounding factors should be ignored. Of course the process is easier when the number of variables is small, with few confounds than in more complex systems. This accounts for the fact that the natural sciences started by studying planetary motion, where the only relevant variables were time and position in space. It follows that explanation of complex biological systems *might* be possible; but, if they are, a would-be theoretician needs far more facts at his finger tips, compared to physical systems where classical explanations proved successful. But the reasoning needed may be relatively simple. Mathematical reasoning has a more limited role than in physics; and may be less formal, akin to that in the humanities (that is large-scale scholarship). However, unlike the humanities, there is a firm conclusion to be reached, and the possibility of prediction, and decisive verification or refutation.

To expound this further, I want to give two examples in which I have been involved, both starting in neuroscience and moving to psychiatry, where cross-level explanations have been emerging. This is very much synopsis, and I'll go into more detail later in the morning

----

From instrumental conditioning
to a model of psychosis

The rubric of instrumental conditioning, the discovery of the self-stimulation phenomenon, and the impact of this on understanding psychosis. In the first half of the twentieth century a great deal of work by psychologists in animals and humans analysed learning using associationist paradigms. The rubric of instrumental conditioning, learning by reward and punishment if you like, while far from a complete account of learning in any species, is arguably important in most learning systems, as described psychologically. This is the "higher level" in my example. What could be the lower level, its neurobiological basis? In the early 1950s, James Olds and Peter Milner addressed this issue.

Peter Milner

(Note that Peter Milner, the theoretician of the pair, was trained as an engineer before becoming a psychologist. He was no doubt used to analysing physical systems with built-in feedback loops.) He reasoned that there must be an internal reinforcement system in the brain. Therefore, by linking an animal's behaviour directly to this system (by-passing sensory systems which normal activate it), behaviour could be reinforced, regardless of its usual motivational significance. The reasoning was classic cross-level explanation in the natural philosophy tradition. It led to the celebrated "self-stimulation", or "brain stimulus reward" paper, published in 1954. Behaving rats, with electrodes implanted in their brain, were able to lever-press to deliver electric pulses to their brain. With some electrode placements, the rats repetitively stimulated their brains, regardless of other prevailing drives. From this, a vast body of research appeared, on electrophysiological, pharmacological and anatomical aspects of the internal reinforcement system. In due course this had major influence in psychiatry, as evidence accrued that a major part of this reinforcement system involved pathways linking the midbrain with the forebrain using the messenger substance dopamine, and it was also realised that antipsychotic drugs were dopamine antagonists.

I became aware of this literature in the early 1970s, and used the idea to explain a singular paradox about effects of antipsychotic drugs: While they block dopamine receptors quickly, the beneficial effects accumulate over weeks or even months. From that insight the idea grew that psychosis was an exaggeration of dopamine's reinforcement function, expressed through distinctively human cognitive processes rather than through outward behaviour. My first paper on this was published in 1976, and it was very much as an attempt to comprehend what had recently been my own lived experience.

Miller 1976

In the early 1980s Rick Beninger from Queen's University (Kingston Ontario) independently came up with a similar concept, and, since we met in 1989, we have worked together, and published several papers together.

Rick Beninger

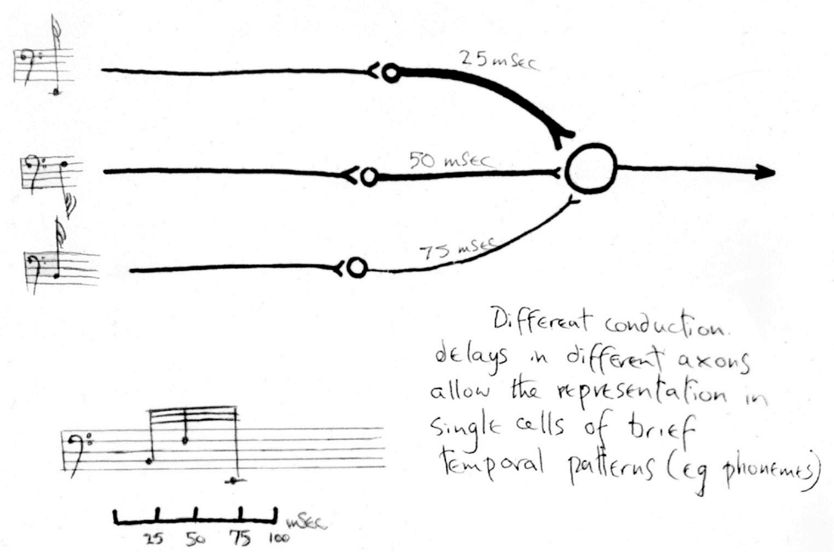
Today this is mainstream understanding of psychosis, although terminology has changed. The idea has practical implications for how antipsychotic drugs should best be prescribed, perhaps leading to more rational prescription than at present, and even better medications.

This first example was about active psychotic states in the disorder called schizophrenia. Psychosis is a state of mental turmoil, which, with modern drug treatment, is usual transient. The second example is about the enduring *non-psychotic* psychological traits, underlying such episodes, which may be present before, during or after periods of psychosis. To link psychological processes to neurobiology, in a truly explanatory way, we obviously need to know a lot about the nerve cells of which the brain is composed.

Schematic neurone

Much *is* known about nerve cell bodies and their biophysics, but so far, it has not proved very useful for cross-level explanations. However, one part of a neurone has been neglected, the humble axon, the “nerve fibre”. We have known the physical basis of action potentials in axons since the 1950s, and in the peripheral nervous system we have known about conduction properties (e.g., conduction velocity) for longer than that. In the brain, evidence on axonal conduction velocity is scanty. As a post-doctoral student in Oxford in the early 1970s I was recording from single cortical neurones in the cerebral cortex of anaesthetised cats, and obtained data on the range of conduction velocities in populations of axons connecting different parts of the cortex. Some axons had conduction times far longer than anyone would have been guessed at the time. The experiment certainly was biased *against* detecting neurones with such slow-conducting axons. Bearing in mind these likely biases, and scaling things up to brains the size of humans, it is likely that different axons in a typical pathway connecting parts of the cerebral cortex together have conduction times (from cell body to synapse) ranging from a few msecs (in rapidly-conducting axons) up to a few hundred msecs (in slowly-conducting ones). The latter have conduction times long enough to be relevant for psychological processes in intact humans. Thus we are within reach of cross-level explanations of psychological findings in terms of neuronal structure and function. I first used this concept to explain a psychological finding in 1981: Each consonant speech sound <give example> is, in acoustic terms, a succession of acoustic events occurring in sequence over a period of about 100 msec.

Paula Tallal figure



Suppose that the connections are sufficiently rich that a typical neurone in the area where speech sounds are recognised has inputs from the primary acoustic region, with a wide variety of conduction times. One might then imagine then that each axon is specific to a particular sound frequency. Thus *this* axon, with conduction time of 25 msec would be specific for <whistle>, *this* one with time of 50 msec for <whistle>, and *this* one with conduction time of 75 msec of <whistle>. Then for a sequence like ***this*** <whistle> all signals would converge on the right-hand neurone at the same instant, the neurone would fire, and the neurone would have the capacity to recognise the brief pattern. Of course acoustic events in a consonant speech sound occur too fast for us to be *aware* of the sequence, but the principle is the same.

Now, we know that human perception of consonants is done better with the left than the right hemisphere. I suggested that the left hemisphere has a richer repertoire of "long axonal delay lines" (which, in a population of axons, represent patterns spread over intervals up to 100 msec) than the right (where conduction hypothetically is faster, the hemisphere then being better for analysing *instantaneous* patterns). By the 1990s this idea had been worked up to produce a broader theory of cerebral asymmetry, the central hypothesis being that, in most pathways, the left hemisphere has axons whose spread of conduction times was longer intervals than in the right. So my central hypothesis for normal cerebral asymmetry was as follows:

Two central hypotheses

From there, I went on to explore an idea which had been around since the late 1960s, that there was *abnormal* laterality in schizophrenia. In the end I could account for many non-psychotic traits of schizophrenia, in terms of the hypothesis that, overall, there is a relative lack of rapidly conducting axons in schizophrenia, these being replaced by slowly-conducting ones. My magnum opus on this was published in 2008.

Miller, 2008

Anyone who wants to grasp the full range of psychological functions to be explained in terms of population-distributions of axonal conduction times, should look at three of my books.

Three books using ACT as premise

The reasoning in these works, as well as the assumptions, have yet to be given a proper critique, and there are empirical implications yet to be tested. I await these things with great interest.

Messages

There are several lessons here: (a) Cross-level explanation in the brain and behavioural sciences *is* possible, including ones related to mental disorders (where they may lead to strategies for treatment). (b) When the explanations work, so can concepts of mental illness become validated, based on proper scientific rationality. Those new concepts may cut across traditional concepts (such as concepts of disease) established in less rational fashion; or they may bring together concepts which were hitherto separated. (c) Those examples given employ just two principles from neuroscience. There are many other such principles, which may be crucial for other explanations. (d) Framing cross-level explanations is impossible unless *someone* is familiar with information at both the "upper level" (i.e. psychological findings, behaviour, symptoms or first-person accounts of experience) and the lower level (details of various aspects of brain biology). That means either that neuroscientists need education about fine details of mental disorders, or that psychiatrists, and clinical psychologists need education about details of neuroscience. This may require change in education in both areas. (e) There is no algorithm by which one can say which upper level details are to be linked with which lower level facts to form an explanation. Evidence at both levels is more complex than in physical systems, so would-be theoreticians need to be extremely well-read, before they have a chance to find the links. This, I think, is more important than skills in simulation or mathematical analysis (although those methods may sometimes be useful).

Objections

People object here that we do not have enough evidence to launch large-scale, library-based theoretical research in psychiatry. I disagree profoundly. It is *not* that we do not have enough information, but rather that too few people know enough of what *is* known to make progress, have no confidence that explanations can be found (because there are no traditions for this), and have no idea how to proceed. Information that has accumulated over the last 100 years at both the higher and lower levels is overwhelming, staggering, much of it (not all) solid empirical data, if only we knew how to use it; but who reads it? Who tries to assimilate it even at one level, let alone across levels, or over different fields? Very few individuals make it their job to know enough to make cross-level explanation really work. That is an issue for how scientists' careers are conceived and administered, rather than for what is possible, were those constraints to be relaxed.

What I suspect has happened is that a new technique is discovered, experimenters rush to exploit it; good papers are published, some not so good; but when the whole area is reviewed it is too messy and complicated to make sense of it, and no-one knows enough in other fields or at other levels to do it. Then another technique is invented, so researchers, always looking to publish good papers and please their masters rather than achieve understanding, switch to the new area; and they do it again. . . and again . . . and again. Each time there is vast expense, and profits for those who make equipment, but little progress in concepts or in understanding. The image which comes to my mind is of a vast orchard, stretching over the horizon. Wherever one looks, one sees trees hanging low, overburdened with ripe fruit, ready for picking. Tending this orchard obviously took prodigious labours of dedicated gardeners in times gone by; yet no-one picks the fruit, and few know the existence of this orchard, and its enormous potential; and yet, since the fruit (all those research papers) are securely archived, the fruit will not become over-ripe and fall from the trees.

Everlasting orchard

Career structure of would-be theoreticians

What this means is that a would-be theoretician should not expect to *combine* theoretical work with experimental or clinical work (both of which demand complete commitment). S/he needs to be a dedicated theoretician. We need a new breed of theoretician-scholars, respected by and respectful of experimental disciplines, both knowing there is mutual benefit to come from the other's approaches, one looking for predictions to test, the other making predictions which the other *can* test. Both understand that (as in physics) good experimenters, and good theoreticians are different sorts of people, with different habits of thought, not to be evaluated on a single scale for "research assessment". In the natural philosophy tradition this has grown ever since the time of Copernicus. It is now desperately needed if the brain and behavioural sciences are to make progress. If it could be brought about, in my view, progress would go further, be faster, be more secure, and would be *much cheaper.*

Conclusions:

Let me draw some general conclusions.

(i) After Newton (in the eighteenth century), in Britain and Europe, we had what is called "the Age of Reason". Today we have what might be called the "Age of Evidence". My view is that reasoning without evidence and evidence without reasoning are equally stupid. We need both, intimately connected, just as Francis Bacon advised.

(ii) The issues in psychiatry *are* fundamental, more so than elsewhere in medicine. In the seventeenth century, apart from the struggle to define what "natural science" could be, an undercurrent of deeper debate, dealt with essentially metaphysical issues, about the meaning for words like "nature", and "causation", and the relation between religious notions of the time and emerging natural science. Should scholars continue using Aristotle's ideas about "final cause", or was there a better notion of causation ("antecedent cause" as we now call it). In psychiatry today, there are also underlying metaphysical questions, about the relation between mind and brain (or equivalently between the subjective and objective worlds), and the nature of causation, issues about determinism etc. The germ theory of infectious disease, or theories of autoimmune disease, neoplasia etc, required no such original thinking at the metaphysical level.

(iii) At a practical community level, concepts of mental disorder defined for scientific purposes do not have the same status as diagnoses used in clinical practice, although they may be *related* to diagnoses. Psychiatry should recognize the unique personality of each person as a central concept, more than does general medicine. So, diagnosis in psychiatry may have a different role from that in general medicine. In addition, diagnoses may need to be adjusted pragmatically to each society, whereas scientific concepts do not.

(iv) Of course, what I have been saying also has implications for research administration, research assessment exercises, funding of research, styles of scientific publication, and for career structure of academic scientists. I think we could perhaps leave those matters for later.

Let me end with words of Francis Bacon. In *"The New Atlantis",* published posthumously in 1626 he used the situation of the "great unknowns" in distant places in the Age of Discovery to develop ideas on a possible better society. Now we don't know the location of this society. It is *assumed* to be somewhere in the western Atlantic; but, on the basis of meticulous analysis of his text, I have reached the conclusion that it was actually somewhere on the Southern coast of Australia, somewhere not far from here; so their voyage had clearly gone *much* further off course than hitherto believed. Anyway, in his concluding pages the author describes the academy in that society, including details on deployment of labour therein:

We have three that try new experiments, such as themselves think good. These are called Pioneers or Miners. We have three that draw the experiments of the former into titles and tables, to give better light for the drawing of observations and axioms out of them. These we call Compilers. We have three that bend themselves, looking into the experiments of their fellows, and cast about how to draw out of them things of use and practice for men's life. . . Then, after divers meetings and consults of our whole number, to consider of the former labours and collections, we have three that take care out of them to direct new experiments, of a higher light, more penetrating into Nature than the former. These we call Lamps. We have three others that do execute the experiments so directed, and report them. These we call Inoculators

----

This glimpse of Bacon's futuristic vision reveals key principles: Freedom of enquiry for researchers; interdependent roles of theory and experiment; decisions taken interactively rather than by top-down control; predictions to test hypotheses; and even what is now “technology transfer”.

Finally, an invitation: If anyone here wants to join me in the enterprise of building truly explanatory theories in areas of mental disorder in which they are interested, ones from which a more solid system of classification can be derived, I would like to hear from you.

Invitation: Contact details
