

## Psychiatry's early breakdown and the rise of the DSM

On a chilly Wednesday morning in late January, I pass through the gates of my university after a fraught drive through London's rush hour. With two minutes left on the clock, I make my way hurriedly to the ground floor of the lecture theatre. Today I am expected to deliver my first lecture on critical psychiatry. As I enter the room it feels more close and cramped than usual, as nearly every student on the course has decided to attend (which, I must add, doesn't always happen on cold January mornings). The students are preoccupied as I approached the lectern and start quietly ordering my notes. Many of them are chatting intently, some are tapping on laptops or mobiles, while a few eager souls (in the front row, of course) quietly sit waiting for me to begin.

'Right everyone, settle down, I have a great piece of research I want you to consider. You'll like this one, trust me, so please listen closely.' I clear my throat and begin.

Some years ago during a balmy April, a group of eight academics conducted a dramatic experiment, months in preparation. As part of the experiment they individually presented themselves at different psychiatric hospitals dotted around the United States. Each academic then told the psychiatrist on duty they were hearing a voice in their head

that said the word 'thud'. That was the only lie they would tell; otherwise, from that point on they would behave and respond completely normally. All of them were admitted into their respective hospitals. And all were diagnosed with serious mental disorders and given powerful antipsychotic pills. All the while they acted completely normally. The experimenters thought they would be in for a couple of days and then be discharged, but they were wrong. Most were held for weeks, and some in excess of two months. They could not convince the doctors they were sane. And telling the doctors about the experiment only compounded the problem. So it quickly became clear that the only way out was to agree that they were insane, and then pretend to be getting better.

Once the leader of the experiment, Dr David Rosenhan, got out and reported what had happened, there was uproar in the psychiatric establishment. Rosenhan and his colleagues were accused of deceit. One major hospital challenged Rosenhan to send some more fake patients to them, guaranteeing that they would spot them this time. Rosenhan agreed, and after a month the hospital proudly announced to the national media that they had discovered 41 fakes. Rosenhan then revealed that he had sent no one to the hospital at all.\*

For a moment there is stunned silence in the lecture room, quickly followed by some chuckling and surprised chatter. I now have their full attention. Three or four hands shoot up.

---

\* Here I paraphrase from Adam Curtis' brilliant BBC documentary, *The Trap* (2007).

‘Hold your questions for now everyone. I’ve another series of experiments to tell you about first. These occurred around the same time as Rosenhan’s experiment, and were equally devastating for psychiatry.’

These experiments explored the following question: ‘Would two different psychiatrists diagnose the same patient in the same way?’ To answer this, the researchers presented the same set of patients to different psychiatrists in different places, to see whether their diagnoses would match up. When the results came in, the situation did not look good. Taken en masse, they revealed that two psychiatrists would give different diagnoses to the same patient between 32 and 42 per cent of the time.<sup>1</sup> And this troubling result was confirmed by another series of studies showing that psychiatrists in the United States and in Russia were twice as likely to diagnose their patients as schizophrenic as their colleagues in Britain and Europe.<sup>2</sup> This meant that the diagnosis you could be assigned not only often depended on who your psychiatrist was, but on where your psychiatrist was located. How could you therefore trust your diagnosis, when a different psychiatrist was likely to diagnose you with something else?

I told my students about these experiments, because in the history of psychiatry they were considered game-changers. They plunged psychiatry into severe crisis in the 1970s by exposing that there was something terribly wrong with the diagnostic system. Psychiatrists were not only defining sane people as insane, but when two psychiatrists at any given time were faced with the same patient, they would assign different diagnoses nearly half the time. So why were these critical mistakes being made?

The profession was desperate for an answer. And when one finally emerged, the course of psychiatry would be altered for good. It turned out there was a serious problem with the centrepiece of the entire profession, the psychiatrist's bible – the DSM.



So what, you may ask, is the DSM? To answer this question, please follow me into the office of Dr Herbert Pardes, one of America's leading psychiatrists. To give you some idea of his professional standing, just consider his CV. He was former chair of Columbia University's Department of Psychiatry (the most powerful psychiatry department on the globe); former president of the American Psychiatric Association (the more glitzy US equivalent of the Royal College of Psychiatrists), and finally, former director of the largest psychiatric research organisation internationally (the National Institute of Mental Health). In short, if there were a CEO of psychiatry, then Herbert Pardes was probably it.

Pardes welcomed me into his office with an easy smile and a warm handshake, 'I'm glad we've finally managed to make this meeting happen,' said Pardes kindly. 'Come on over, take a seat.'

Once Pardes and I had settled comfortably in his unexpectedly grand office, the first topic I pressed him on was the DSM. 'If you don't understand the history of the DSM,' insisted Pardes, 'you cannot hope to understand modern psychiatry.' The DSM is shorthand for the *Diagnostic and Statistical Manual of Mental Disorders* and is the book listing all the psychiatric disorders that psychiatrists believe to exist. 'So the DSM contains every mental disorder with

which you or I could be potentially diagnosed', said Pardes, 'and that's its significance.'

Pardes then briefly recalled the DSM's journey from its modest 130 pages in 1952 to the 886 pages it boasts today. In short, the first edition of the DSM was written in order to solve a problem that had plagued the profession for decades. Until the 1950s, psychiatrists working in different places possessed no shared dictionary in which all the disorders were clearly defined and that carefully listed each disorder's core symptoms. Without this dictionary, the behaviour that one psychiatrist called 'melancholic' or 'depressive' another psychiatrist was likely to call something else. So this made communication between psychiatrists in different places almost impossible.<sup>3</sup> 'If I say to another psychiatrist that I have tried the drug Thorazine on 250 people with paranoid schizophrenia,' explained Pardes, 'what happens if this other psychiatrist's definition of paranoid schizophrenia is not the same as mine? Well, our discussion becomes meaningless. So the DSM was developed to try to identify and standardise the symptoms characteristic of any given mental illness – anxiety disorder, phobia, mood disorder and so on.' Every psychiatrist was then expected to learn this list so that different psychiatrists in different places would all be working from the same page.

Once the first DSM arrived in the 1950s, psychiatrists were expected to use the dictionary in the same standardised way still in operation today. For instance, if you go and visit a psychiatrist tomorrow because you're feeling down, the psychiatrist will ask you to describe your symptoms. The purpose of this is to try to work out from your symptoms what diagnosis from the dictionary you should be

assigned. For example, if you report feeling tense, irritable and panicky, and that you have been feeling this way for over two weeks, then you are likely to be diagnosed with one of the anxiety disorders. Whereas if you mention that you're feeling sad, teary and lethargic and are experiencing disrupted sleep, then you are more likely to be diagnosed with one of the depressive disorders. Of course, sometimes your symptoms will not fall neatly into any single category, but rather span two or three. In this case your problem will be considered 'comorbid' – namely, that you are suffering from a disorder that is occurring simultaneously with another (perhaps you suffer from major depression as well as panic disorder). But whether your condition is comorbid or not, the diagnostic process is the same – your psychiatrist attempts to match your symptoms as closely as possible to one of the diagnostic labels listed in the book.

Now here comes the problem. And it's a problem that still afflicts psychiatry today. How does your psychiatrist know if he or she has assigned the correct diagnosis? Is there a safe and reliable way that he or she can test, objectively speaking, whether the diagnosis given is the right one? I put this question to Pardes: 'Well, one way to test whether the diagnosis is correct is to apply a scientific or biological test [such as a blood, urine or saliva test] or some other form of physical examination to assess, firstly, whether a patient has a mental disorder, and, if so, precisely what disorder they suffer from. But the crucial problem for psychiatry is that we still have no such objective biological tests.'

In other words, unlike in other areas of medicine where a doctor can conduct a blood or urine test to determine whether they have reached the correct diagnosis, in

psychiatry no such methods exist. And they don't exist, as Pardes also intimated, because psychiatry has yet to identify any clear biological causes for most of the disorders in the DSM (this is a pivotal point that I'll talk about more fully in coming chapters). So the only method available to psychiatrists is what we could call the 'matching method': match the symptoms the patient reports to the relevant diagnosis in the book.

These facts, although at first glance appearing innocuous, are crucial for understanding why psychiatry, in the 1970s, fell into serious crisis. They help us explain why psychiatrists were not only guilty of branding sane people as insane (as the Rosenhan experiment revealed), but also guilty of regularly failing to agree on what diagnosis to assign a given patient (as the 'diagnostic reliability' experiments showed). Psychiatry was making these errors because it possessed no objective way of testing whether a person was mentally disordered, and if so, precisely what disorder they were suffering from. Without such objective tests, the diagnosis that a psychiatrist would assign could be influenced by their subjective preferences, and as different psychiatrists were swayed by different subjective factors, it was understandable that they regularly disagreed about what diagnosis to give. This is why these early experiments were so dramatic for the profession: they produced for the first time clear evidence that psychiatric diagnosis was at best imprecise, and at worst a kind of professional guesswork. And so without any objective way of testing the validity of a diagnosis, psychiatry was in peril of falling far behind the diagnostic achievements of other branches of medicine.

A solution was needed, and fast.



Under the leadership of the American Psychiatric Association (APA), the profession in the 1970s plumped for a radical solution. It decided to tear up the existing edition of the DSM (then called *DSM-II*) and start again. The bold idea was to write an entirely new manual that would solve all the problems beleaguering *DSM-II*. This new manual would be called *DSM-III*, and its central aim would be to improve the reliability of psychiatric diagnosis and thereby answer the mounting criticisms that were threatening to shatter the profession's legitimacy.\*

The first step the APA took was to set about finding someone to lead the writing of *DSM-III*. The APA needed a person highly competent, energetic and daring, but also someone who had experience with psychiatric classification. After sifting through countless candidates and enduring many frustrations, the APA finally settled on a man called Dr Robert Spitzer, who was based at Columbia University's medical school. Spitzer had been a young and up-coming psychiatrist when the earlier *DSM-II* had been written, and he had also been minimally involved in that project. But most importantly, he appeared to have the drive and vigour needed to get the job done. The APA was sufficiently impressed with his qualities, so they hired him in 1974 to start work on *DSM-III*. Little did Spitzer know at

---

\* I was often told that poor diagnostic reliability was not the only driver for the DSM's reform. There was also a need to match DSM terminology to that used in the *International Classification of Diseases* (ICD). However, Robert Spitzer, Melvin Sabshin and other leaders in the APA knew that the reliability issue was paramount and that the DSM must make that issue its priority.



the time that his appointment as Chair of *DSM-III* would ultimately make him the most influential psychiatrist of the 20th century.

The first thing Spitzer did to reform the DSM was to assemble a team of fifteen psychiatrists to help him write the new manual. This team was called the DSM Taskforce, and Spitzer was its outright leader. So in the mid-1970s the Taskforce set about writing a kind of New Testament for psychiatry: a book that aspired to improve the uniformity and reliability of psychiatric diagnosis in the wake of all its previous failings. If this sounds all very intrepid, well, that's pretty much what it was. Spitzer's Taskforce promised a new deal for psychiatry, and there was a lot of pressure on them to deliver.

So what precisely did Spitzer do to try to set things right? How was he going to make psychiatric diagnosis more reliable and scientific? His answer was simple. The DSM needed to be altered in three major ways:

- Many existing disorders would be deleted from *DSM-II*.
- The definitions of each disorder in the old DSM would be expanded and made more specific for *DSM-III*.
- A new checklist would be developed for *DSM-III* to improve the reliability of diagnosis.

Let's briefly look at each of these alterations more closely. The first involved Spitzer deleting some of the more unpopular and controversial mental disorders. These included some of the disorders introduced into psychiatry by psychoanalysis, a discipline with important differences from

psychiatry (see footnote below).<sup>\*</sup> In the 1970s psychoanalysis had fallen out of vogue in psychiatry, along with many disorders it had introduced to the previous DSM. One of the most controversial of these was homosexuality. Indeed, in the *DSM-II* homosexuality was listed as a mental disease. It was described as a 'sexual deviation' and was located in the same category as paedophilia.<sup>4</sup> While some psychiatrists felt it was wrong to brand homosexuality an illness, the main push to remove the disorder largely came from outside pressure groups including the gay rights movement. These groups asked why a normal and natural human sexual preference had been included in the DSM as a mental disease, especially when there was no scientific evidence to justify its inclusion. Surely it was prejudice rather than science that had placed homosexuality on the list?

Many psychiatrists were not so sure, but the APA, perhaps sensing the change in public mood, decided to consult the wider psychiatric community for their views. So at the APA convention in 1973 all the attending members were asked to vote on what they believed: was homosexuality a mental disorder or not? The vote was closer than expected:

---

\* What is the difference between psychoanalysis and psychiatry? Or between psychiatry, psychology and psychotherapy, for that matter? It can be summarised this way: a psychologist researches different aspects of our mental lives – cognition, memory, perception, etc. They are not clinicians, unless they have undertaken a specialist postgraduate training in clinical psychology or psychotherapy (the 'talking cure'). The psychotherapist or psychoanalyst, on the other hand, has trained at the postgraduate level to treat patients with the 'talking cure' – they do not have to be medical doctors (and so do not prescribe medications). Psychiatrists are medical doctors who have later specialised in psychiatry. Some psychiatrists practise one form of psychotherapy or another but most do not, nor do they have to. Today, most psychiatrists diagnose disorders and prescribe and monitor medications.

5,854 psychiatrists voted to take homosexuality out of the DSM, while 3,810 voted to keep it in. And because the 'outers' were in the majority, homosexuality ceased to be a mental disorder in 1974 and was therefore not included in Spitzer's *DSM-III*. It was politics and not science that had removed the disorder from this list. As we continue, it's worth holding that thought in mind.

To turn now to Spitzer's second alteration, this involved making the definitions of each mental disorder more specific and detailed. The idea was that if each disorder could be defined more precisely, psychiatrists would be less likely to misunderstand the disorders and therefore misapply them to patients. The problem with the earlier *DSM-II*, Spitzer had argued, was that its definitions of disorders were too open to interpretation. So, for example, in *DSM-II* 'depressive neurosis' was defined in a single sentence: 'This disorder is manifested by an excessive reaction of depression due to an internal conflict or to an identifiable event such as the loss of a love object or cherished possession.'<sup>5</sup> Spitzer believed that such vague definitions explained why psychiatrists regularly gave different diagnoses to the same patient. If a word in the dictionary were poorly defined, people would not know how to use it properly. The same was the case with psychiatric diagnoses. This imprecision was why, as Spitzer said, for *DSM-II*, 'there are no diagnostic categories for which reliability [is] uniformly high ... [and why] the level of reliability is no better than fair for psychosis and schizophrenia and is poor for the remaining categories.'<sup>6</sup> Spitzer's hope was that by sharpening the definitions there would be less scope for personal interpretation, which in turn would mean diagnostic reliability would rise.

Finally, to help improve diagnostic reliability further, Spitzer's team created criteria for each disorder that a patient had to meet in order to warrant the diagnosis. So while, for example, there are multiple symptoms associated with depression, it was somehow decided that a patient would need to have at least *five* of them for a period of at least *two* weeks to qualify for receiving the diagnosis of depression. The only problem was: on what grounds did Spitzer's team decide that if you have five symptoms for two weeks you suffered from a depressive disorder? Why didn't they choose six symptoms for three weeks or three symptoms for five weeks? What was the science that justified putting the line where Spitzer's team chose to draw it? In an interview in 2010, the psychiatrist Daniel Carlat asked Spitzer this very question:

*Carlat:* How did you decide on five criteria as being your minimum threshold for depression?

*Spitzer:* It was just consensus. We would ask clinicians and researchers, 'How many symptoms do you think patients ought to have before you would give them the diagnosis of depression?,' and we came up with the arbitrary number of five.

*Carlat:* But why did you choose five and not four? Or why didn't you choose six?

*Spitzer:* Because four just seemed like not enough. And six seemed like too much [Spitzer smiles mischievously].

*Carlat:* But weren't there any studies done to establish the threshold?

*Spitzer:* We did reviews of the literature, and in some cases we received funding from NIMH to do field trials ... [However] when you do field trials in depression and other disorders, there is no sharp dividing line where you can confidently say, 'This is the perfect number of symptoms needed to make a diagnosis' ... It would be nice if we had a biological gold standard, but that doesn't exist, because we don't understand the neurobiology of depression.<sup>7</sup>

I expect that by now some of you may be scratching your heads. Wasn't the whole point of Spitzer's reform to make psychiatric diagnosis a little more scientifically rigorous? But what, you may ask, is rigorous about a committee drawing arbitrary lines between mental disorder and normality? And what is scientific about asking the psychiatric community to vote on whether existing disorders should be removed from the DSM? In other words, in the name of making psychiatric diagnosis more scientific, had Spitzer's team continued to make use of the unscientific procedures that had dogged the construction of earlier manuals?

As important as this question is, I'll refrain from answering it right now, because there is a more crucial question to be addressed first: did Spitzer's reforms actually work? Did they solve the reliability problem? I mean, if you went to see two different psychiatrists independently today, would they be likely to both assign you the same diagnosis?

In an interview for *The New Yorker* in 2005, a journalist called Alix Spiegel asked Spitzer that very question. His answer was unequivocal: 'To say that we've solved the reliability problem is just not true', said Spitzer. 'It's been improved. But if you're in a situation with a general clinician

it's certainly not very good. There's still a real problem, and it's not clear how to solve the problem.<sup>8</sup> Here Spitzer admits something that many within the profession agree with: diagnostic reliability, despite the reforms, is still woefully low.

According to a study published in the journal *Psychiatry* in 2007, for instance, which asked a group of psychiatrists whether they thought psychiatric diagnosis was now reliable, a full 86 per cent said that reliability was still poor.<sup>9</sup> It was not only their clinical experience that led them to this conclusion, but also presumably their familiarity with existing research, including work undertaken by Spitzer himself to find out whether his reforms had worked. Its conclusions were not reassuring. For example, you'll remember that I said before Spitzer's *DSM-III* two psychiatrists would give different diagnoses to the same patient 32 per cent to 42 per cent of the time. Well, Spitzer found that after his reforms psychiatrists were now disagreeing around 33 to 46 per cent of the time – results indicating the very opposite of diagnostic improvement.\* And these disappointing figures are consistent with other more recent studies also implying that reliability is still poor. For example, another study published in 2006 showed that reliability actually has not improved in 30 years.<sup>10</sup>



---

\* The psychologist Paula J. Caplan argues that one study showed that when different psychiatrists were diagnosing patients from the Axis II group of disorders (basically the personality and developmental disorders) their diagnoses were the same only about two-thirds of the time (66 per cent). Whereas for the remaining disorders they were the same only about half the time (54 per cent). See: Caplan, P.J. (1995), *They Say You're Crazy*. New York: Da Capo (pp. 197–200).

An obvious question for the British reader is whether poor diagnostic reliability is a problem in the UK? After all, in the UK we have alongside the DSM the *International Classification of Diseases* (the ICD). Perhaps the ICD leads to greater reliability than the DSM? Although this is a reasonable question to ask, when we take the research en masse, it actually shows that using the ICD leads to no greater diagnostic reliability than using the DSM.<sup>11</sup> This may partly explain why in countries like Britain where the ICD is used along with the DSM, many mental health researchers and professionals often prefer the DSM.<sup>12</sup> In fact, the National Institute for Clinical Excellence (the body that sets the clinical guidelines for the NHS) now recommends the use of the DSM over the ICD for disorders including depression.<sup>13</sup> Also, in my own experience of working in the NHS, the DSM is a very influential manual. But even if you wanted to dispute its precise impact, and as an article in the *British Journal of Psychiatry* put it: ‘we’d still not avoid all the problems that beset the DSM [here in Britain]. Both manuals were developed and classify mental disorders in pretty much the same way. As the DSM writes: “the many consultations between the developers of the *DSM-IV* and the *ICD-10* ... were enormously useful in increasing the congruence and reducing meaningless differences in working between the two systems”.’<sup>14</sup> Herbert Pardes also confirmed this to me when recounting that ‘the DSM worked very closely with the ICD to get worldwide cooperation between diagnostic categories’. In other words, diagnostic reliability is a problem for international psychiatry – whichever manual you employ, the reliability rates are broadly the same.

This leads me to one final point about the reliability problem that would be perilous to overlook: what would happen if some day reliability rates in psychiatry were to improve dramatically? This question is important because it reveals a more fundamental problem for psychiatry that it has yet to solve: even if every psychiatrist on the globe independently diagnosed the same patient with the same disorder (for example, with ‘social anxiety disorder’), this would still not *prove* that social anxiety disorder actually exists in nature, that it’s actually a discrete, identifiable biological disease or malfunction of the brain. You require much more than mere agreement to prove that. You need hard evidence. Unless our sciences can test whether what we agree on is objectively the case, agreement counts for nothing from a scientific standpoint. So even if psychiatrists reach high diagnostic agreement at some future point, this would not prove that the mental disorders with which they diagnose patients actually exist as valid disease entities. There need to be other procedures to establish that. So the issue is: are there other procedures? And if so, what exactly are they?

This question is so central to the entire psychiatric enterprise that I decided to ask Robert Spitzer myself.