Psychological Medicine, 2005, **35**, 467–474. © 2005 Cambridge University Press
DOI: 10.1017/S0033291704004210 Printed in the United Kingdom

INVITED REVIEW

Beyond major depression

GORDON PARKER*

School of Psychiatry, University of New South Wales, and Black Dog Institute, Prince of Wales Hospital, Sydney, Australia

ABSTRACT

Background. The DSM concept of 'major depression' has accrued increased status but demonstrated limited utility since inception.

Method. An historical overview of models of depression and the initially perceived advantages presented by the concept of 'major depression' are presented before detailing its limitations in application.

Results. Challenges to the utility of 'major depression' are provided by examining its conceptual model, its validity, its utility and the limited information generated in aetiological and treatment efficacy studies.

Conclusion. It is argued that the concept of 'major depression' has led to sterility in depression research and clinical practice, and that there is a need for a paradigm shift in modelling and classifying the depressive disorders.

> 'Taxonomy is described sometimes as a science and sometimes as an art, but really it's a battleground.' (Bryson, 2003).

THE PROPOSITION

Hodge (1971) described the whiplash injury as 'Magnificent in its simplicity and how it seizes the imagination of patients, doctors and lawyers'. The same can be said about the term 'Major Depression'. I will argue that the concept, initially understandable and even admirable, has accrued entity status and explanatory properties beyond its station. Designed principally to address a general problem of diagnostic reliability, it has subsequently failed to demonstrate diagnostic validity and utility. More importantly, its transubstantiation to 'entity' status has led to conceptual confusion, and to sterility in both depression research and clinical management. Further, its reification has largely extinguished consideration of alternative classificatory models for depressive disorders. To understand its origins and initial positioning, there is a need to overview some historical developments.

* Address for correspondence: Professor Gordon Parker, School of Psychiatry, University of New South Wales, and Black Dog Institute, Prince of Wales Hospital, Randwick, 2031, Sydney, Australia.
(Email: g.parker@unsw.edu.au)

HISTORICAL BACKGROUND

As reviewed previously (Boyce & Hadzi-Pavlovic, 1996), a two-class 'binary model' for classifying the depressive disorders had a lengthy history, with 'endogenous/psychotic' *versus* 'neurotic/reactive' dichotomous labels early in the twentieth century respectively weighting biological and psychosocial determinants. A contrasting 'unitarian' model (i.e. there being one depressive condition, varying along a 'dimension' or 'continuum' of severity) emerged for several reasons. In North America, the Myerian dimensional model of psychological illness

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

influenced depression conceptualization, with his biopsychosocial model not easily accommodating illness 'categories'. In the UK, opposition to the binary model came from Mapother (1926), who argued against the prevailing Kraepelin model, and held that manic-depressive psychosis was a quantitative disturbance aligned along a dimension that also included the neuroses. Further, arguing that as distinctions between 'psychotic' and 'neurotic' expressions of depression had no aetiological, natural history or treatment implications, Mapother concluded that it was pointless to attempt to distinguish them. The subsequent introduction of multivariate statistical techniques led to several decades of attempts to determine whether 'depression' was best modelled as a unitarian construct or according to the binary model. Boyce & Hadzi-Pavlovic (1996) suggested that, in the absence of a 'knock-down argument' by the time DSM-III committees addressed their tasks, the 'only sensible verdict' in regard to the binary model was the Scottish one – *not proven*.

Thus, lacking firm empirical support for a binary model and in the absence of any accepted alternative categorical model, it was not inappropriate for the revisionist DSM-III classificatory system to implicitly favour a dimensional model for the depressive disorders, contrasting and defining 'major' and 'minor' disorders by symptom number and duration parameters – and building on the RDC prototype (Williams & Spitzer, 1982). Fifth-digit coding also effectively allowed depression subclassification of more putatively categorical disorders (i.e. psychotic and melancholic depression), although the DSM-III authors left the modelling status of the latter as open questions. For example, the DSM-III Introduction notes that 'The inclusion of Major Depression With and Without Melancholia as separate categories in DSM-III is justified by the clinical usefulness of the distinction', however this did not 'imply a resolution of the controversy as to whether or not these conditions are in fact quantitatively or qualitatively different' (p. 6).

Several DSM-III processes and components shaped the concept of 'major depression'. First, its introduction in 1980 occurred in a political context. As noted by one of America's leading psychiatrists, 'the development of DSM-III represents a fateful point in the history of the American psychiatric profession … a significant reaffirmation on the part of American psychiatry to its medical identity and its commitment to scientific medicine' (Klerman, 1984). It was a development that sought to redress the dominance of psychoanalysis and psychodynamic theorizing. It sought to bring Psychiatry back to its medical roots by deriving explicit diagnostic criteria, so generating a reliable and valid classificatory system, and one which asserted 'how knowledge about mental disorders should be organized' (Kirk & Kutchins, 1992). If, as it has been held, that to understand depression is to understand Psychiatry, it was important for the DSM-III committee to take no risk on depression classification and to have a system that allowed diagnoses to be made with some consistency – reliability was to be prioritized.

Second, and as noted in its Introduction (pp. 7–8), DSM-III was 'descriptive'. The actual DSM-III description (pp. 210–211) of a 'Major Depressive Episode' (or MDE) included vegetative and endogeneity symptoms, severe psychomotor disturbance, and worthlessness and guilt ranging up to delusional intensity – a descriptive profile that, when read, circumscribes the 'more biological' depressive disorders to a striking degree. However, one of the guiding principles to DSM-III was to have clinical criteria 'described at the lowest order of inference necessary to describe the characteristic features of the disorder' (p. 7), risking less constrained boundaries, and thus, both allowing inclusion of the 'less biological' disorders and even inclusion of less clear-cut 'cases'. So while the *descriptive profile* was of a disorder of such severity that one would be reluctant to assign such patients to any research study that delayed immediate treatment, if not hospitalization, the *diagnostic criteria* bar for many items was set very low, risking 'over-diagnosing' major depression in practice. The same concern holds for the requirement in subsequent DSM manuals for the individual to be impaired. The DSM-IV Criterion C requires that the symptoms of a major depressive episode 'cause clinically significant distress or impairment in social, occupational, or other areas of functioning'. The first component would seem self-evident while the first '*or*' actually allows 'distress' to be sufficient and effectively obviates any

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

mandatory impairment requirement. Even if the latter is respected, its vague definition again risks the inclusion bar being set too low. Such issues also pose an intrinsic risk to reliability estimates.

Third, DSM-III was largely atheoretical – as inclusion of aetiological theories was judged to present an obstacle to 'clinicians of varying theoretical orientations'. Its atheoretical genesis, however, risked begetting limited downstream explanatory power. That general operational principle almost certainly limited subsequent consideration of a key reality – that unless the classificatory system is underpinned by a valid explanatory model for the disparate depressive disorders, that system will be intrinsically flawed.

While the criterion set has varied somewhat in terms of number and content of actual descriptors, 'major depression' has now been in existence for a quarter of a century. While it has long outgrown its initial aspirations, they are worth acknowledging. As for other DSM diagnoses, it provided a 'standard frame of reference' (Kendell & Jablensky, 2003), promoted diagnostic agreement and communication, and thus advanced clinical discourse and allowed research studies to be replicated.

CHALLENGES TO THE CONCEPT

However, how reliable is a diagnosis of 'major depression'? Post-implementation studies rapidly raised concerns. Thus, in one USA study (Anthony *et al.* 1985), comparing DSM-III diagnoses made by psychiatrists and by lay interviewers using a standardized interview schedule, the kappa coefficient was 0·25 for major depression, with the researchers identifying a large number of possible determinants. While Kirk & Kutchins (1992) have mounted a more general challenge to the claim that 'DSM-III resolved the reliability problem' (p. 15), it must be recognized that such identified limitations may not be specific to DSM-III or to any other classificatory system, and more reflect difficulties intrinsic to the field of assessing non-categorical psychopathological symptoms. It may be that the definition of 'major depression' is as good as it gets in terms of allowing the prevalence of 'significant depression' or 'clinical depression' to be estimated, but it cannot be concluded that DSM-defined 'major depression' is, of necessity, an intrinsically reliable diagnosis. If reliability is important to the research endeavour, it needs to be demonstrated within the particular study rather than assumed.

While major depression is commonly viewed as a 'valid' psychiatric diagnosis, it fails to meet any of the orthodox criteria for validity. It does not have a clear-cut clinical picture (a depressed mood state being the only obligatory component), its boundaries are unclear (because they reflect dimensional rather than natural cleavages), its natural and treatment history are difficult to predict at the individual level, while cause and response to treatment are again more related to factors in the individual sufferer rather than being integral to the disorder.

In terms of diagnostic utility, Kendell & Jablensky (2003) suggest that 'utility' emerges if the diagnosis provides (in part) 'nontrivial information about prognosis and likely treatment outcomes'. In comparison to diagnostic validity, this property is central if any diagnostic entity is to claim its place in the sun. I will argue that major depression fails the test of providing meaningful information about aetiology, prognosis and treatment.

Let us consider aetiology first. By definition, DSM avoids aetiological considerations, while clinicians accept that multiple psychological, social and biological factors may contribute to episodes of major depression. Their relevance is determined more by sample selection (and representation within *that* sample of differing biological or psychosocial causes) than by anything integral to major depression *per se*. Thus, the construct has no aetiological specificity. Does it allow identification of any specific pathophysiology? In one review, Hickie (1996) concluded that studies of patients with major depression had 'failed to demonstrate any coherent pattern or neurobiological changes' or 'replicate key biological correlates across different research groups, age cohorts and treatment settings'. Such results largely reflect the heterogeneity intrinsic to major depression, compounded by the heterogeneity provided by differing samples.

Prognosis? If aggregated data are examined, an account can be provided for major depression, both in terms of natural and treated

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

history. Average length of episode, likelihood of recurrence, chance of developing a chronic disorder, rates of disability and chance of suicide can all be estimated – even with some consistency across studies. However, major depression has no intrinsic prognostic explanatory power when applied at the individual level, limiting its utility to the clinician.

Treatment specificity is perhaps the most important component to diagnostic utility. However, and as reviewed elsewhere (Parker, 2004), the randomized controlled trials (RCTs) of treatments for major depression suggest similar overall efficacy levels (of some 50–55%) for most treatments – including old antidepressants, new antidepressants, manualized psychotherapies and St John's Wort – an equipotency result that invites the Dodo bird verdict (Holmes, 2002) of 'Everyone has won, and all must have prizes.' The additional failure of many antidepressants to differentiate from placebo in studies of major depression – half, in one analysis of FDA studies (Khan *et al.* 2002) – also challenges the diagnostic construct. If the RCT reference category of 'major depression' results in problematic, non-specific and relatively meaningless information for clinical application, then its utility must be queried. Further, the in-built design component of having diagnostic criteria calibrated to the 'lowest order of inference' presumably contributes to RCT subjects being likely to have less severe and more spontaneously resolving disorders, thereby clouding demonstration of any true specificity to differing treatments.

It is not surprising that a non-specific diagnostic category will contribute to non-differential treatment outcomes. If dyspnoea were the non-specific diagnostic category, and comparator treatments of bronchodilators and antibiotics assessed, it might be anticipated that there would be minimal differentiation of those quite widely varying treatments at the population level, with differentiation entirely dependent on the underlying representation of asthma and chest infection. Further, if the majority of dyspnoic subjects had asthma, and the trial involved a placebo-controlled study of an antibiotic, then a truly effective treatment (for asthma) might emerge as ineffective.

Thus, on logical grounds alone, we should not expect RCTs of interventions for major depression to be particularly informative. Moreover, rather than interpret such non-specific findings as reflecting methodological and application limitations to the criterion diagnosis of major depression and RCT procedures, results tend to be interpreted at face value. For example, in one published guide (Duvorsky & Duvorsky, 2002), we read that mild and moderate single episodes of major depression 'can be treated with antidepressants or psychotherapy' (p. 219) and that 'all antidepressants currently available are equally effective' (p. 220). In Treatment Guidelines published by my professional college (RANZCP Guidelines, 2004), it is stated that moderate to severe depression can be treated by four differing drug classes and two differing psychotherapies. Such an 'All roads lead to Rome' non-specific treatment model hardly inspires but, more importantly, is it clinically credible? Worse, does it advance clinical management? In practice, the equipotency model (allowing all principal therapies to be viewed as comparably efficacious) promotes professionals treating patients according to their preferred treatment model. Thus, the individual patient is likely to be 'fitted' to the therapist's treatment paradigm rather than the therapy being 'fitted' to the specific determinants or characteristics of the disorder (or disorder 'type').

CURRENT STATUS

Despite lacking clinical meaning or treatment utility, DSM-defined 'major depression' has become the 'biggest game in town'. In terms of cachet value, 'major depression' impresses the media, judges and insurance companies, and may ensure hospitalization and medical insurance coverage. In contrast to the unaesthetic and consonant-driven term 'dysthymia', it sounds major league and assists destigmatization by its gravitas. Like 'whiplash injury' it is also magnificent in its simplicity and has been majestic in its impact. It has become 'clinical depression'. It circumscribes a range of heterogeneous conditions, homogenizes them, and has come to be viewed as an entity, a construct that has been misconstrued. 'Major depression', the media may ask, 'What causes *it*?' Similarly, clinicians are encouraged to treat '*it*' in certain ways, while researchers pursue '*it*' in studying

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

causes and treatments. As the reference diagnosis, candidate treatments are required to demonstrate their capacity to treat *it* – and are thus tested as if they have universal application in RCTs. Like the proverbial elephant examined by blind men, it allows quite disparate conclusions. Many educational campaigns – both to health professionals and to the public – explicate a model proclaiming major depression as a disease, reflecting chemical changes in the brain and therefore requiring antidepressant medication. Conversely, others equally cogently argue that it reflects dysfunctional cognitive schema that require cognitive behaviour therapy, while others reason persuasively that it is a consequence of interpersonal distress and requires interpersonal psychotherapy. It has fooled some of us for some of the time, but it now risks fooling us all of the time.

Not only has major depression become the biggest game in town, it has become the 'only game'. DSM-III is not to be blamed here. As noted earlier, it posited diagnostic subtypes such as melancholic depression, allowing for aetiological and treatment differentiation studies to explore their utility along with the utility of major depression. But such 'subtypes' became marginalized, being viewed as having little intrinsic explanatory power or treatment ramifications. For instance, Dubovsky & Dubovsky (2002) noted that, while the Dexamethasone Suppression Test had the capacity to distinguish melancholic and non-melancholic depression, 'this is not of great practical importance, given that both types of depression are treated similarly'. Leaving aside the capacity of DSM decision rules to validly identify 'melancholia' and the literature suggesting distinct treatment differences for melancholic and non-melancholic depression, the point is that major depression not only subsumes depressive subtypes by definition but, by its hegemony, it has subverted any interest in a subtyping model.

In terms of its diagnostic utility, a diagnosis of 'major depression' is no more explanatory than – pursuing the analogy – 'clinically dyspnoic'. In both instances, a clinical domain is identified, inviting a second stage prior to any prescribed treatment – the search for the underlying cause and application of a rational treatment – rational in addressing symptom reduction and/or addressing the underlying cause. The dyspnoic individual may be breathless because they have overly exerted themselves in a fun run, and the clinician might then work to an expectation of spontaneous remission after resting (i.e. an adjustment disorder). Alternately, the dyspnoic patient may have asthma, pneumonia or a pulmonary embolus, encouraging quite differing treatment strategies. By contrast, the diagnosis of major depression tends to be viewed as sufficient in and of itself.

How to proceed? Rather than continue to view and reify major depression as a diagnostic entity with explanatory power, or to argue for its demise, there is wisdom in regarding 'major depression' as a first-stage estimate of the likelihood of 'clinical depression', inviting a second stage involving more fine-focused diagnostic subtyping. For that, we need to recognize that 'depression' can exist as a disease, a disorder, a syndrome, and/or be a normal or abnormal reaction to salient stressors. To address that reality, a single dimensional or single categorical model is unlikely to be appropriate. As argued here, dimensional or continuum models (and the same concerns expressed here for DSM 'major depression' hold for the dimensional – but marginalized – ICD-10 model of 'severe', 'moderate' and 'mild' depression) have not advanced depression research or management in application A simple dichotomous binary model is implausible, unless one is a heterogeneous residue after any other and more definitive type is identified. Assuming a 'mix and match' modelling paradigm, the need then is to identify those depressive disorders that are best modelled categorically and those for modelling dimensionally. For the latter, there is a need to identify the components (i.e. severity or other parameter) that might best be dimensionalized.

IS PSYCHIATRY ALONE?

Psychiatry is not alone in having difficulty in deriving a precise taxonomy for the depressive disorders. Bryson (2003) has detailed how pre-Linnaeus, Botany had a 'highly whimsical' classification system, with animals categorized along qualitative (e.g. terrestrial or aquatic) and quantitative (e.g. large or small) parameters, while post-Linnaeus, phyla estimates range from the twenties to the eighties, partly dependent

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

on whether the biologists were 'lumpers' or 'splitters'. Again, Bryson (p. 389) informs us about Palaeontology, where hominid types had risen to more than 100 by the 1950s, where a binary model was proposed in the 1960s (*Australopithecus* and *Homo*), and where now 'some twenty types of hominid are recognized in the literature today. Unfortunately, almost no two experts recognize the same twenty', and 'The only way a name becomes accepted is by consensus, and there is often very little of that' (p. 389).

Such a 'lumping and splitting' process is clearly evident in examining the classification of the depressive disorders over time – phenomenological description (particularly by European psychiatrists) generating more and more putative subtypes, followed by a reductionist move to a binary (e.g. 'endogenous *versus* reactive') typology, before 'major depression' largely obviated pursuit of such complexities. Depression subtyping now risks being viewed as of little importance, sterile, or – even worse – not even requiring consideration. However, if great minds are able to be applied to downstream fine-focused domains such as genetics and neuroimaging, why not have such concentrated application given to upstream model building? The aphorism 'What's the use for running if you're on the wrong road?' is salient to current applied research.

QUO VADIS?

In the absence of diagnostic markers, the task involved in modelling 'depression' may be more akin to interpretive anthropology where respect for 'thick description' is given, and where Geertz (1975) has suggested that its progress as a science 'is marked less by a perfection of consensus than by a refinement of debate'. Let me risk advancing a 'mix and match' (i.e. categorical and dimensional) model as a template for consideration.

First, let us assume that there are certain depressive disease 'categories' (e.g. psychotic depression and/or melancholia) that were judged as worthy of including in DSM-III and, as noted by Paykel (2002), are 'still worth keeping'. If categorical, then we need markers that allow both their clinical definition and their boundaries from non-melancholic disorders to be drawn with evident independence. Our research (Parker & Hadzi-Pavlovic, 1996) has argued for the utility of operationally defined and clinician-rated psychomotor disturbance, but there are many other markers that should be comparatively tested in pathophysiological and treatment differentiation studies.

The residual 'disorder' states – which are likely to be intrinsically inter-dependent rather than independent – are more problematical to model (Parker & Manicavasagar, in press). They can, of course, be modelled along DSM and ICD parameters, involving mood-state severity, persistence and recurrency dimensions. They can also be modelled by approximating to clinical formulations reflecting contrasting contributions of stressful events and underlying personality predilections (akin to earlier concepts such as 'reactive depression' and 'neurotic depression'). Other modelling options include weighting clinical features (historically leading to concepts such as 'anxious' and 'irritable/hostile' depressions), triggering or precipitating factors (historically so weakly differentiating as to lead to the demise of the term 'endogenous depression'), or weighting their presumed predisposing or aetiological factors (e.g. 'involutionary' and 'post-natal' depression, and depressions with a significant personality disorder contribution) and akin to dementia classification.

As each of the determining domains and the end-point states lack specificity *when tested alone*, and any integral disorders will, of necessity, be inter-dependent to some degree, we should again consider the anthropological paradigm of defining 'fuzzy sets' and allowing definition to occur across several domains or axes. As any integral disorders are unlikely to be purely driven by precipitating or by predisposing factors, a diathesis-stress aetiological model might also be usefully tested. If valid, in the sense that causes link with disorder configuration, then the 'fuzzy set' needs to accommodate descriptors that operate at multiple levels. The only depressive condition currently so positioned in DSM-IV is 'Atypical depression' – with the long-standing suggestion of treatment response specificity (i.e. to MAOI medication) – and which can be viewed (Parker *et al.* 2002) as having a primary personality (i.e. interpersonal sensitivity rejection) base associated

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

with certain clinical features (e.g. hypersomnia, hyperphagia). This model suggests a paradigm (Parker, 2000; Parker & Manicavasagar, in press), where the search might be for identification of 'spectrum' non-melancholic depressive disorders linking predisposing personality style and clinical phenotype.

The possibility, then, of pursuing personality-weighted endophenotypes is intuitively attractive, although, as defined by Gottesman & Gould (2003), this would limit consideration to heritable constructs, a limitation when some predisposing personality styles may more be the consequences of early and profound deprivational experiences. A weighting then to personality, as against genetically driven temperament, would appear favoured.

First *et al*. (2004) suggest that the 'holy grail' of clinical utility would be to demonstrate the positive effect of a change in the diagnostic system on clinical outcome. Since the introduction of DSM-III-defined 'major depression' the field has gone backwards in that regard. As noted by Carroll (1989), the antidepressant efficacy of imipramine was established using the '"old-fashioned" diagnostic systems of ICD-8 and DSM-II' while studies using 'major depression' as a criterion fail to demonstrate the superiority of imipramine over placebo treatment. That reality – of non-specific treatment findings when the criterion diagnosis is 'major depression' – has continued for far too long. There is a need to develop candidate subtyping models and then compare their individual capacity to demonstrate differential treatment outcomes against current continuum models.

CONCLUSION

Psychiatry research and practice would benefit from moving beyond 'major depression'. It is neither an intrinsically reliable nor valid entity, and, as for utility, it has failed to provide the field with necessary clarification about aetiology or differential treatment response. Its historical pedigree is worthy of respect (i.e. as a first-pass criterion for the likelihood of depression being of some clinical magnitude) but such a 'diagnosis' cannot provide the template required to advance depression research and treatment.

ACKNOWLEDGEMENTS

Preparation of this paper was assisted by an NHMRC Program Grant and a grant from the Centre for Mental Health, NSW Department of Health. Editorial assistance by Kerrie Eyers is gratefully acknowledged.

DECLARATION OF INTEREST

None.

REFERENCES

Anthony, J. C., Folstein, M., Romanoski, A. J., VonKorff, M. R., Nestadt, G. R., Chahal, R., Merchant, A., Brown, H., Shapiro, S., Kramer, M. & Gruenberg, E. M. (1985). Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis. *Archives of General Psychiatry* **42**, 667–675.

Boyce, P. & Hadzi-Pavlovic, D. (1996). Issues in classification: I. Some historical aspects. In *Melancholia: A Disorder of Movement and Mood* (ed. G. Parker and D. Hadzi-Pavlovic), pp. 9–19. Cambridge University Press: New York.

Bryson, B. (2003). *A Short History of Nearly Everything*. Doubleday: Australia.

Carroll, B. J. (1989). Diagnostic validity and laboratory studies: rules of the game. In *The Validity of Psychiatry Diagnosis* (ed. L. N. Robins and J. A. Barrett), pp. 229–245. APA Series, Raven Press: New York.

Dubovsky, S. L. & Dubovsky, A. N. (2002). *Concise Guide to Mood Disorders*. American Psychiatric Publishing Inc: Washington, DC.

First, M. B., Pincus, H. A., Levine, J. B., Williams, J. B. W., Ustun, B. & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnosis. *American Journal of Psychiatry* **161**, 946–954.

Geertz, C. (1975). *The Interpretation of Cultures*. Hutchinson: London.

Gottesman, I. I. & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intensions. *American Journal of Psychiatry* **160**, 636–645.

Hickie, I. (1996). Issues in classification: utilising behavioural constructs in melancholia research. In *Melancholia: A Disorder of Movement and Mood. A Phenomenological and Neurobiological Review* (ed. G. Parker and D. Hadzi-Pavlovic), pp. 38–56. Cambridge University Press: New York.

Hodge, J. R. (1971). The whiplash neurosis. *Psychosomatics* **12**, 245–249.

Holmes, J. (2002). All you need is cognitive behaviour therapy. *British Medical Journal* **324**, 288–294.

Kendell, R. & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry* **160**, 4–12.

Khan, A., Khan, S. & Brown, W. A. (2002). Are placebo controls necessary to test new antidepressants and anxiolytics? *International Journal of Neuropsycharmacology* **5**, 193–197.

Kirk, S. A. & Kutchins, H. (1992). *The Selling of DSM: The Rhetoric of Science in Psychiatry*. Aldine de Gruyter: Hawthorne, New York.

Klerman, G. L. (1984). The advantages of DSM-III. *American Journal of Psychiatry* **141**, 539–542.

Mapother, E. (1926). Manic-depressive psychosis. *British Medical Journal* **2**, 872–877.

Parker, G. (2000). Classifying depression: should paradigms lost be regained? *American Journal of Psychiatry* **157**, 1195–1203.

Parker, G. (2004). Evaluating treatments for the mood disorders: time for the evidence to get real. *Australian & New Zealand Journal of Psychiatry* **38**, 408–414.

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press

Parker, G. & Hadzi-Pavlovic, D. (1996). *Melancholia: A Disorder of Movement and Mood. A Phenomenological and Neurobiological Review*. Cambridge University Press: New York.

Parker, G. & Manicavasagar, V. (in press). *Modelling and Managing the Depressive Disorders: A Clinical Guide*. Cambridge University Press, Cambridge.

Parker, G., Roy, K., Mitchell, P., Wilhelm, K., Malhi, G. & Hadzi-Pavlovic, D. (2002). Atypical depression: a reappraisal. *American Journal of Psychiatry* **159**, 1470–1479.

Paykel, E. S. (2002). Mood disorders: review of current diagnostic systems. *Psychopathology* **35**, 94–100.

RANZCP Clinical Practice Guidelines Team for Depression (2004). Australian and New Zealand clinical practice guidelines for the treatment of depression. *Australian and New Zealand Journal of Psychiatry* **38**, 389–407.

Williams, J. B. W. & Spitzer, R. L. (1982). Research Diagnostic Criteria and DSM-III. An annotated comparison. *Archives of General Psychiatry* **39**, 1283–1289.

https://doi.org/10.1017/S0033291704004210 Published online by Cambridge University Press