4 CAUSES AND CORRELATES

> Cause and effect in the mental disorders isn't a simple matter. Asking *why* someone is depressed, for example, may be relevant if you're dealing with depression following a post-traumatic stress disorder, because it's useful for planning treatment. But what happened before an episode of manic-depressive illness may be quite irrelevant: the cause lies elsewhere, at a different level of brain organization.
>
> –W.H. Calvin & G.A. Ojemann, *Conversations with Neil's brain* (1994)

Prologue

As I said in the last chapter, I am not a dualist: I do not believe that 'mind' in its fundamental constitution is any different from 'body' (though of course it is experientially). It follows that depression must be a physical disease. Therefore its immediate cause, what directly underlies the *experience* of the disease, is brain dysfunction. This may be both on the chemical level (disordered chemistry) and the structural (loss of or damage to brain tissue). These mechanisms have been empirically observed.

To unpack this further: judging from the way antidepressants and other drugs affect depression, there is clearly a neurochemical component. Judging from the way depressions can arise as reactions to life events there is clearly an element of stress-response and susceptibility to trauma, hence an environmental component. And judging from the way mood disorders frequently cluster in families, and are distributed over pairs of twins, there is a significant genetic component as well. Therefore there will not be a single 'cause' for mood disorder in general or any particular occurrence, but a multitude of causes (and underlying vulnerabilities) at different levels and located at different points in a victim's history, and involving different brain systems. But in the end, regardless of cause, current brain dysfunction is the basis of the depressive experience. This is what depression *is*, and this is where the apparently most successful treatment strategies are aimed, either directly as in drug treatment or indirectly as in psychotherapy. (For the argument that psychotherapy is also a physical treatment of brain dysfunction see chapter 5.)

On causes

A causal theory of an illness is the best foundation for treatment. It provides a framework for directed rather than trial-and-error research, and allows theory-based treatments to be devised and subjected—along with the theory—to rigorous testing. If you think that HIV infection is due to failure to pray to your ancestors, serious testing might well show that no matter how much you pray, unprotected sex with a carrier is still likely to infect you, and all the prayer in the world will not increase your CD4 count. But medicine does not live up to its ideals any more than other professions. It is common to develop successful treatments or preventives for illnesses whose causes are unknown, using medications whose effects are not understood. In the late 18th century, Edward Jenner capitalised on the uninterpreted observation that dairy-maids who had

contracted the relatively mild disease cowpox tended to be immune to the far more deadly smallpox. In 1796 he infected young James Phipps with material from a cowpox lesion, and performed the first (publicised) successful medical inoculation. This began a tradition of lifesaving vaccination, long before the germ theory of disease, almost a century before anyone even suspected the existence of viruses, and long before the existence of the immune system, let alone its complexity, had been imagined.

The history of medicine is full of these happy accidents, many of them still cornerstones of treatment, and still not understood. Clinical trials directed by theoretical knowledge or good hypotheses are clearly better than faith or random trial-and-error, or even than long-term empirical observation, though that often contributes massively.[1] But sadly, the more complex and subtle a disease process is, the less theoretical grasp we usually have of it. 'Cause' in any case is an equivocal and difficult notion; and psychiatric illness is so complicated, and our knowledge of the brain and how it interacts with the environment so imperfect, that here causation has a special obscurity. More of our treatments are products of serendipity than we would like. But we will need some idea of the conceptual intricacy of the notion 'cause of an illness', if we are to understand the point of the longstanding and often acrimonious debate on the causes of depression, and its treatment—or even how it can be treated at all.

Let us begin with a much simpler kind of disease: what causes malaria? The answer in antiquity, and until the discoveries of Sir Ronald Ross in the late 1890s, was 'swamps'—or their noxious emanations (Italian *mal' aria* 'bad air'). The evidence was good: the Romans knew that draining swamps and shutting windows at night to keep out the bad air reduced the incidence of malaria.. Now we would say that malaria is caused by a parasite of the genus *Plasmodium* transmitted by the bite of an infected *Anopheles* mosquito. The ancients were really keeping mosquitoes out of their houses by closing their windows at night, and preventing successful breeding by draining swamps. Their procedures worked, but not for the reasons they imagined.

But in fact they were right: swamps *do* cause malaria; but they were less directly right than we are when we say that plasmodia cause malaria via infected mosquitoes. But what does it mean to say that 'plasmodia cause malaria'? The parasite itself does not cause the disease; what we call malaria is simply a byproduct of its activities. Baby plasmodia get into red blood cells and mature, and then burst out of them; the destruction of red blood cells and the chemicals released when they burst cause the periodic chills, fever and malaise that we define as 'the disease'.

Mosquitoes need standing water to lay their eggs in; these hatch into larvae; the survivors pupate and turn into adults; the females feed on the blood of mammals or birds; and one who bites an infected host ingests plasmodia which mature inside her, and enter her salivary glands.

---

[1] Of course all the error in the world will not mark a trial as unsuccessful if there is a strong enough prior belief, and a technique for making the errors go away. People still pray to their ancestors or other beings for cures, dip all or part of themselves in magic waters, wear talismans, etc. With this mindset, 5000 unsuccessful trials are nothing; if by chance a cure and an act coincide, that is good enough, a 'proof'. This is what distinguishes superstition or magic from science.

The next time she bites somebody she passes them on; and the result of their elaborate life-history in the body is the syndrome we call 'malaria'. This is not even a particularly complicated example; it merely shows that many interesting phenomena result not from a simple juxtaposition of 'causes' and 'effects', but from *chains* of interrelated events and preconditions.

Philosophers distinguish two kinds of linear causes : *ultimate* and *proximate.* An ultimate cause is the beginning of a causal chain; a proximate cause is the nearest to the effect in question —*given the current limitations of our knowledge.* There are however degrees of proximalness and ultimalness. We still have not asked why bursting of red blood cells causes these particular symptoms, so even the apparently proximate 'bursting of red cells' is not fully so. For practical purposes we may stop pursuing the 'absolutely' proximate and concentrate on what we do have, contenting ourselves with good working approximations, elliptical causal stories. It still makes sense to say that plasmodial infection causes malaria, or that the bite of an infected *Anopheles* causes malaria. The 'right answer' depends on whether the matter at hand is treating patients or controlling mosquitoes. Causality, ultimate or proximate, can be interpreted at different levels of resolution.

This however just scratches the surface. Printed across the top of a packet of cigarettes on my desk is a prominent health warning saying 'Danger: Cigarettes cause cancer'. Is this true? Well, yes and no. This causal claim looks as if it were a 'law of nature', e.g. a statement like 'loss of the body's entire blood volume causes death'. But in ordinary untidy language-use it is not meant that way. What *is* true is that smoking hugely increases the risk of certain cancers; you are at least 60% more likely to get lung cancer if you smoke than if you do not, and something like 90% of lung-cancer victims are or have been smokers. But increase in actuarial risk is not the same thing as causation; you have no way of knowing whether or not *you* will get cancer, only that your risk is higher. Many smokers do not get it, and a good number of non-smokers do.

The health warning is not a causal claim at all: it is an admonition about gambling. The real message is that betting against the House is likely to fail in the long term. Read critically, it says only that under the appropriate (unspecified) conditions you have a particular (unspecified) risk of getting cancer. This is not a condemnation of the claim, rather of its (fairly typical) use of 'cause'. We will see that there are elements in the story of depression that are actuarial in the same sense—the conditions are still not fully specified or even understood, but we have a good picture of what is actually happening and what is predisposing.

Similarly, an expression like '*Mycobacterium tuberculosis* causes TB' still needs a more subtle interpretation. You can be exposed to the bacillus and not get infected (your immune system can dispose of it); or you can be asymptomatically infected because your immune system makes antibodies quickly enough to prevent the disease from taking hold, though the bacteria are still hiding inside you; or you can come down with symptomatic TB. Causes are not, conceptually, only ultimate or proximate; we must also distinguish between *necessary* and *sufficient* conditions. The presence of *M. tuberculosis* in the body is a necessary but not sufficient condition for getting TB. Smoking cigarettes is neither a necessary nor sufficient

condition for getting lung cancer, though it is strongly predisposing. The loss of one's entire blood-volume is a sufficient condition for death, but not a necessary one. There are many ways to die—though all of them will finally involve your heart stopping. This is the one cause that is both necessary and sufficient.

Good science, as Sir Peter Medawar once said, is 'the art of the soluble' (1967). Good medicine might be 'the art of the usable'. Often we have to make do with quite imperfect approximations to both.

'Psychological causation': two classical theory-types

I will now look briefly at two purely 'psychological' accounts of the causes of depression. I do not think either of them tells anything near the whole story, but they are worth considering because of their high historical profiles, and their influence on current thinking and therapeutic practice. Both also make claims that contain elements of significant truth, and even when they are most erroneous or empty, that in itself can make us think about things we might otherwise have taken for granted.

Historicity

Much contemporary discourse about depression and other psychiatric illnesses is driven by a loosely 'Freudian' model of mental architecture and ultimate causality. Whatever psychotherapists have to deal with is a matter of the patient's personal history, and wherever it comes from, it currently inhabits The Unconscious. Depression is the result of 'repression', either of the memory of traumatic or unpalatable events or losses, or of unpalatable fantasies or 'unresolved conflicts'. In some neo-Freudian frameworks the problem is 'fixation', failing to mature past certain psychosexual landmarks, lack of proper attachment to a parental figure or defective nurturing. In Freud's later work it is anger—which ought to be directed against some Other but for certain reasons cannot be—turned against oneself. Depression is caused by events in one's personal, post-natal past that have left disguised traces, or by inappropriate strategies for dealing with the results of one's history. The therapist's task is to undo this damage by bringing hidden contents to light, to 'reconstruct' the psyche, to help the patient confront and neutralise the demons of the past.

A later development of this view, coming from the 'biological psychiatry' rather than the psychoanalytic tradition, derives depression not from repressed memories or conflicts, but at least partly from environmentally imposed and internalised trauma or stress. Childhood events, inside or outside the family, dysfunctional relationships with parents or other caregivers, abuse, often apparently 'forgotten', set the scene for later depressive episodes or chronic disorders. But not all depressions, whether 'reactive' or 'endogenous', acute or chronic, derive from *old* historical sources. Current stressors can have the same effects. A middle-aged person with no prior history of mood disorder may plunge into a shattering and potentially lethal depression

after the loss of a loved one, the breakup of a marriage or being retired.

Perhaps confusingly, these recent stressors produce an illness indistinguishable from the recognised forms of chronic depression— those that do not seem to be responses to some local, recent life-experience but originate in the distant past, or even apparently nowhere. The patient's symptoms give no information about what sort of event (if any) triggered the depression, or when. And most important, the *subjective* quality of the depressive experience appears to have no relation to the recency or nature of whatever might have induced it. This stress-response perspective is more narrowly historical, though the history may be temporally shallow. I will return to this later, as it is one of the most important causal factors in mood disorder. But not in precisely the way it is outlined above.

Psychic archaeology and psychodynamics: the Freudian legacy[2]

Sigmund Freud's brilliant speculative geography of the human mind has become a cultural commonplace; it still dominates our unreflecting knowledge, our stock of metaphors. This is true even if we have never read a word he wrote. In some ways this is unfortunate. Freud was a great poet, mythmaker and dramatist, and at times a supremely gifted clinical observer. The popular conception of mental structure has been irrevocably—and in some particulars correctly—changed by his fleshing out of a hidden world of autonomous mental activity that we are not conscious of. But whatever his cultural importance, he was not (except in his early days as a laboratory neurologist) a scientist in the conventional sense.[3] The Freudian model of the psyche and its activities is essentially a work of art or prophecy. Psychoanalysis, like traditional healing or homeopathy, sometimes appears to work;[4] but (a) there is no respectable theory behind it; (b) worse, there is no way of setting up controlled statistical tests that can distinguish genuine efficacy from placebo effect or natural remission; and (c) worst of all, in many respects it does not cohere with the rest of our knowledge of the world.

Freud created a rich and powerful mythology, in which the various parts of the psyche, dealing in the currency of an occult something called 'psychic energy', are personified and

---

[2] For useful if brief accounts of Freudian and post-Freudian theory see the articles 'Freud' and 'Freudianism: later developments' and their cross-references in Gregory 1987. For a devastating critique of Freudian theory and practice, richly illustrated with autobiographical and other examples, see Sutherland 1998.

[3] According to one influential scholarly tradition, it is not even clear that Freud was always honest about reporting his own results, but seems to have cooked some of his data or even invented it to prove theoretical points. For a detailed and immensely readable study in this vein (with rejoinders from opponents), see Crews 1997.

[4] Whether there is even a demonstrable placebo effect in psychoanalysis is problematic, as 'cure' and 'improvement' are ill-defined. See Medawar 1967, and Fieve 1997: chapter 11.

engage in titanic battles.[5] The Ego, the Id, the Superego are locked in perpetual warfare; the analyst 'treats' the patient by forcing him to re-remember this (arguably imaginary) battle in his past, and through this understanding or re-remembering to demystify, come to terms with, and eventually tame the great mythical figures whose battles have left psychic scars—or at least get them to declare a truce. The same can be said of many of Freud's followers: Jung in particular was as much a shaman and mythographer as Freud, and his great archetypal characters, Anima, Animus, Shadow, etc. are also participants in epic mythological conflicts.

Probably the most influential psychoanalytic treatment of depression is Freud's famous essay, 'Mourning and melancholia' (1917). This is a typically Freudian mixture of insight and nonsense. The core (and true) insight is that melancholia (unipolar depression, or the depressive side of bipolar disorder) and mourning for lost 'significant others' share many features. Anthony Clare characterises these well as 'a painful sense of despair over loss, a significant lack of interest in the outside world, a loss of the capacity to love and a marked reduction in activity' (1994: 91). Freud notes however that in melancholia there is—compared to ordinary mourning—'self-reproach and an irrational expectation of punishment', as well as a vagueness about what the loss being (*ex hypothesi*) mourned actually is. According to Freud, the loss in melancholia is internal: 'In grief, the world has become poor and empty; in melancholia it is the ego itself'.[6]

He gets the description mostly right (though mourning too is often accompanied by overwhelming guilt), but needs a mythology to explain it. The well observed parallel with ordinary mourning, once transformed into a cause, requires a loss that could be unconsciously mourned. Freud finds it in a (hypothetical) childhood or other relationship tainted by some failure of the loved one. This causes the emotion originally invested in the loved one to be projected onto the image that remains as a player in the unconscious drama, and the patient then identifies this internal character with himself:

> If one listens patiently to a melancholic's many and various self-accusations, one cannot in the end avoid the impression that often the most violent of them are hardly applicable to the patient himself, but that with insignificant modifications they do fit someone else, someone whom the patient loves or has loved or should love [...] The woman who loudly pities her husband for being tied to such an incapable wife as herself is really accusing her husband of being incapable, in whatever sense she may mean this. There is no need to be greatly surprised that a few genuine self-reproaches are scattered among those that have been transposed back. These are allowed to obtrude themselves since they help to make the others and make recognition of the true state of affairs impossible.

This is vintage Freud. The analyst knows (because the theory tells him) what the 'true' situation

---

[5] This may sound unfair. After all, the details of neurotransmission and neuronal function were not yet known, and there was little if any way of describing mental function except through metaphor. What is problematic is the scale and 'imperialism' of Freud's metaphorical world, the degree of personalisation, and the unsupported (and untestable) attributions of causal power that he gave to his metaphors. As well as (see below) aspects of his argumentative style.

[6] This and the following quotation taken from Clare 1994*:* 91ff. and Barondes 1998: 34ff.

is, without requiring any specific empirical evidence from the case in question. The (generic) woman is really doing something other than what she appears to be doing, and the analyst comes equipped with this knowledge *a priori*. But there is an escape clause as well. *Some* self reproaches may be just that, but these are unimportant, only serving to mask the 'true state of affairs', so that the patient cannot recognise it; only the analyst can. But if this is the origin of depression, it would seem that manic depression is a problem, as the expressed affect is precisely the opposite. Freud realised this, but did not come up with a unified account until his *New introductory lectures on psychoanalysis* (1933). In manic depression,

> the most remarkable characteristic is the way in which the super-ego— you may call it, but in a whisper, the conscience—treats the ego. The melancholiac during periods of health can, like anyone else, be more or less severe towards himself: but when he has a melancholic attack, his super-ego becomes over-severe, abuses, humiliates, and ill-treats his unfortunate ego, threatens it with the severest punishments, reproaches it for long forgotten actions [...] and behaves as though it had spent the whole interval in amassing complaints and was only waiting for its present increase in strength to bring them forward, and to condemn the ego on their account [...] It is a very remarkable experience to observe morality, which was ostensibly conferred on us by God and planted deep in our hearts, functioning as a periodical phenomenon. For after a certain number of months the whole moral fuss is at an end, the critical voice of the super-ego is silent, the ego is reinstated, and enjoys once more all the rights of man until the next attack. Indeed in many forms of the malady something exactly the reverse takes place during the intervals; the ego finds itself in an ecstatic state of exaltation, it triumphs, as though the super-ego had lost all its power or had become merged with the ego, and this liberated, maniac [sic] ego gives itself up in a really uninhibited fashion, to the satisfaction of all its desires.

Such can't-lose argumentation led Sir Peter Medawar to remark that 'the property that gives psychoanalysis the character of a mythology is its combination of conceptual barrenness with an enormous facility in explanation' (1984: 66). Even though Freudian and neo-Freudian theory no longer dominate the psychiatric mainstream, much of the basic framework, some of the 'actors', and notions like 'hidden content', 'repressed memory' and 'unconscious mind' remain even in more sophisticated and empirically based causal theories, and the role of childhood 'drama'(correctly, though in other, quite non-Freudian ways) is still important.

Depression as 'erroneous cognition'

There are probably as many ideas about the causes of depression as there are theoretical frameworks. Many are either rewordings of each other or of standard wisdom; but at least one, which has become therapeutically important, starts from childhood in a very different way. This is the 'cognitive' framework developed first by the American psychiatrist Aaron Beck, and now incorporated into the increasingly popular praxis called Cognitive Behavioural Therapy (CBT: Beck 1967, Beck *et al*.1979, Scott & Beck 2008).

Beck's most original idea is that depressive affect is the result of 'erroneous' or 'faulty' *thinking*; negative conceptual frameworks or 'constructions of the world' feed back into mood.

This is of course trivially true, but whether it is central to the origins of the depressive process is debatable. Beck's model does however make some novel points and produces a coherent scenario for the origin of depressive states. In summary:[7]

> Early Experience [infancy and early childhood: acquisition of negative attitudes toward self, world, and future]
>
> ⇓
>
> Formation of dysfunctional (unconscious) assumptions [negative affect colours perception and behaviour]
>
> ⇓
>
> Critical incident(s) [some stressor affects the already vulnerable child: loss, abuse, etc.]
>
> ⇓
>
> Assumptions activated [the response to the stressor is in terms of the unconscious assumptions]
>
> ⇓
>
> Negative Automatic Thoughts [unpremeditated negative 'thoughts' concerning either past, present or future: e.g. guilt, conviction of worthlessness or failure]
>
> ⇓
>
> Symptoms of depression (behavioural, motivational, affective, cognitive, somatic)

The central pillar is 'an underlying theoretical rationale that an individual's affect and behaviour are largely determined by the way in which he structures the world' (Beck *et al.*1979). Depressive themes arise not from unconscious conflicts or biochemical dysregulations, but from the 'primary triad': (i) construing experiences in a negative way; (ii) therefore viewing oneself negatively; (iii) therefore viewing the future negatively. On this theoretical basis Beck and his followers constructed a therapy which claims to treat depression by correcting the patient's 'erroneous' or 'faulty' assumptions about himself or the world. A contemporary practitioner gives this exposition of the causal narrative outlined above (Bolon 1998: 25):

> Experience leads people to form assumptions (or schemata) about themselves and the world. These [...] subsequently organise perception and govern and evaluate behaviour, e.g. "I must do well at everything I undertake".
>
> Problems arise when critical incidents occur, which mesh with a person's own system of beliefs, e.g. the belief that personal worth depends entirely on success could lead to depression in the face of failure.
>
> These dysfunctional assumptions produce an upsurge of negative automatic thoughts, e.g. "I am a write off [sic], I will never amount to anything".
>
> They are negative in that they are associated with unpleasant emotions, and automatic in that they pop into people's heads rather than being the product of a deliberate reasoning process. They may be interpretations of current experiences, predictions about future events, or recollections of past things.
>
> These then lead to the symptoms of depression.

This raises a number of questions. How much of our thinking (even if we are not depressed) is

[7] This diagram is based on one in Bolon 1998. Material in [ ] is added commentary.

rational anyway? Doesn't a great deal, perhaps most, of our everyday mental activity consist precisely of thoughts that 'pop into' our heads? How much 'deliberate reasoning' does the average person do in the course of a day, except for very special purposes? Aren't there plenty of people, depressed or not, who hardly ever do any, either through lack of interest or not being bright enough? Is the content of consciousness under 'our' control, in the sense that we are agents who dictate it?

The basic narrative is profoundly unhistorical—as well as unbiological, and in one major way descriptively inadequate. If depressive affect is due to misconstructions of the world, 'cognitivised' as it were out of mood disturbance due to traumatic or self-devaluing experience, how is it that nonhuman animals can suffer the classic signs of depression as well, and be treated effectively with the same drugs that work for people? Would we want to say that a dog becoming anorexic and anxious or huddling in a corner is a victim of his 'automatic negative thoughts', provoked by a poor self-image due to a bad puppyhood? If treatable depression, under conditions of adult stress, occurs in animals as distant from us as dogs, as it commonly does, then surely its origins are likely to lie—both historically and anatomically—much deeper than propositional cognition or ideation.[8] In humans, it certainly is true that negative thinking can feed back to and influence mood (just as positive thinking can); but it would seem unlikely that purely neocortical processes could have much in the way of a long-lasting and 'constructive' effect on subcortically provoked emotional states.

A closer look at the relation between conscious thought and unconscious brain activity is interesting in this connection. In 1981 the neurophysiologist Benjamin Libet made a rather odd discovery about the temporal relation between self-perceived conscious intention and what the brain actually does. This has been a source of satisfaction to some and distress to others, because of the peculiar light it casts on 'our' relation to our brains. To simplify, Libet asked his subjects to perform some voluntary action (e.g. move a finger), and measured the 'readiness potential', the onset of an electrical pattern in the brain usually detectable before any complex action. This is apparently an index of what the brain is doing in relation to what the subject perceives as his intention. The results are neatly summed up by Susan Blackmore (1999: 226):

> If you are a dualist you may think that the decision to act must come first [before the readiness potential]. In fact what Libet found was that the readiness potential began about 550 milliseconds [...] before the action, and the decision to act about 200 milliseconds [...] before the action. In other words, the decision to act was not the starting point—a finding that can seem a little threatening to our sense of self.

This may not appear directly relevant to the question of mood, but it is to the underlying problem of the relation between conscious and unconscious cognition. The frontal cortex (the primary site

---

[8] Prozac is now as much a part of the veterinary medical armoury as of the human. I am grateful to Pieter Human for discussing with me in detail the use of antidepressants in his veterinary practice. I assume that nonhuman animals do not have propositional cognition of the kind we do; this may or may not be correct, but there is no hard evidence to the contrary.

for executive and planning activity) is not the only part of the brain that 'makes decisions', and its decisions may be later than those made elsewhere. In other words, we may make (apparently) 'conscious' decisions before we are conscious of having made them.[9]

Mood is not caused by cognitive activity in the usual sense—though it is often accompanied by it, and certainly influences it. If someone is *asked* why they are feeling awful, they may well be able, if they are articulate enough, to come up with *ad hoc* verbal descriptions of states of mind, simply because the context of inquiry is linguistic. These will naturally be presented as 'thoughts' or 'reasons' (judgements, interpretations, beliefs, evaluations, i.e. the 'erroneous thoughts' that Beck's theory requires). But the brain does not appear to work this way: mood can only be interpreted by neocortex *after the fact*. The frontal lobes cannot have anything very interesting to say about mood until they get some subcortical input to base it on. *It is the body, the brainstem and the limbic system that tell the neocortex what to feel*, or at least give it the primary data.[10]

This idea was beautifully laid out, if in a somewhat speculative way, in a famous paper by the American psychologist William James (1884). In essence, rather than it being the other way round, 'we feel sorry because we cry' (190). According to his hypothesis, 'the emotional brain processes not only resemble the sensorial brain processes, but in very truth *are* nothing but such processes variously combined' (188). James has been vindicated by recent research. There is now a modern argument that both emotions and the bodily phenomena that generate them occur prior to any mood alterations or even perceptions of mood. This is summarised in Damasio (2003: 67-9). This book, by a neurologist and neuroscientist, presents, in a very accessible way, the most sophisticated theory of mood and emotion that I have come across. Damasio's argument is based on the fact that the clinical and neurological evidence shows that in changes of mood, and especially in the onset of depression, the first reactions are in the brainstem and hypothalamus, then body posture and facial expression change, and *then* finally the person becomes conscious of the mood. So it seems that mood is a bottom-up not top-down phenomenon, which would appear to be bad news for the cognitive model. It just happens, as James had pointed out over a century ago, that we are acculturated to give priority to our mental experiences, and consider them to be causal, rather than what are visualised as their bodily 'accompaniments'. And this is reinforced by the speed with which the enchained phenomena occur, which was not available to measurement in James' time.

In addition, the focus on 'thought' I think misses a crucial aspect of the way depression is *experienced* by a great many patients (rather than what they may typically say about it). It is perfectly possible, even usual, for the worst part of a depressive episode not to consist of 'thoughts' at all, not even feelings of unworthiness or guilt. The quintessential depressive

---

[9]These comments may be simplistic. For a detailed discussion of Libet's work and its significance see Dennett 2004: chapter 7.

[10]To be fair, Beck does admit that 'in some cases' low mood may contribute to 'negative thinking', and this may be 'a causal theory', but is not the norm (Scott & Beck 2008: 641).

experience is not cognitive but *existential*: imprisonment in a quality of the world, objectless despair and blackness and pain, inability to feel pleasure, lack of will, lack of energy with no ideational content at all.

Let us say a cognitive therapist manages to convince me intellectually, regardless of what I happen to say and may (at the time) think or feel, that I am not a worthless human being and a total professional failure. That would have no effect whatever on my mood. Even in the depths of a depression I would still 'know' (at one level) how good I am, be able to admire my own work, my accomplishments, recognise my skills and abilities, feel guilty only for things I happen to have done that on reflection I would rather not have. But at least in a well matured depression like mine, no propositional or empirical evidence is binding on my beliefs or can alter my mood. *It is my mood that produces the thoughts*, and I may only half-believe them, or even not believe them at all, but I am depressed nonetheless.[11] Peter Kramer, as an experienced clinical psychiatrist, puts this point nicely (1994: 209f): 'the distinguishing feature of the damaged sense of self is its poor responsiveness to evidence, even to evidence that is cognitively appreciated'.

But despite these theoretical and empirical problems CBT seems often to be therapeutically effective. The literature suggests that it can do a great deal by itself in minor depressions, and can support the work of drugs in more major ones, as well as lengthening the intervals between episodes, and may even work for major depression. As we have seen it is quite possible for a misunderstood or not understood medical technique to be effective. My concern here is not with the therapy but with the theory, since my topic here is causation, not treatment.[12]

The origins of the 'medical model': how drugs tell stories

The first significant breakthroughs in treating depression came in the 1950s, and then only through a series of lucky accidents and the equally lucky presence of scientists who understood their implications. These led to a radically new, if by no means universally accepted, understanding of these illnesses. This new understanding generated, for the first time, some reasonably effective therapies—and with them an approach that potentially brought 'mental illness' within the scope of ordinary laboratory science. It became possible in principle for psychiatry to gain some respectability as a part of 'real' medicine. This novel investigative and clinical paradigm led to an understanding of major aspects of what was thought then to be proximate causation; regardless of the fact that it still does not touch fully proximate causation, it is much closer than before. It was (and remains, I think) the most profitable direction for

[11] For a somewhat similar critique, Sutherland 1998: chapter 22.

[12] Beck has recently changed his mind on some issues and adopted a more holistic and complex view of depression, advocating the inclusion of work on stress and genetics into an enriched cognitive paradigm (see Moram 2008).

research, and ultimately therapy.[13]

The first serendipities involved the side-effects of drugs being used to treat other conditions.[14] In the 1950s a new anti-tuberculosis drug called iproniazid was developed. Aside from killing TB bacilli, it often seemed to be a 'psychic energiser', stimulating patients' appetites and generally making them feel healthier and happier, even to the point of euphoria, than treatment of the disease alone ought to have done.

Another discovery involved the reverse effect. Again in the late 1950s a new antihypertensive, reserpine, was introduced, and one of its major side-effects was depression. So we have two unrelated drugs, devised for unrelated conditions, which seem to be related to depression in opposite ways. One induces it, the other seems either to relieve it or to produce semi-euphoric states. The crux was that both, whatever else they were doing, were known to affect certain neurotransmitter systems—in particular those involving serotonin and noradrenaline. Iproniazid is a monoamine oxidase inhibitor (MAOI): it disables the enzymes that degrade amine neurotransmitters after they have bound to receptors. Thus, aside from its anti-TB action, it incidentally increases free intrasynaptic amines. Reserpine, on the other hand, lowers blood-pressure partly by depleting amines (especially noradrenaline) at presynaptic terminals. This is a nice coincidence: a drug that increases amine concentrations is a mood-elevator, and one that depletes amines can induce depression.

It would be sloppy to argue that just because a drug has chemical effect A and produces physical or psychological effect B, the whole secret of B must be the chemistry A. Heartburn is not caused by antacid deficiency. On the other hand it would be daft to dismiss such a correlation out of hand. These and related observations developed into the 'Bioamine Theory' of depressive illness: the claim that certain neurotransmitters, particularly serotonin and noradrenaline (and nowadays dopamine), are crucially implicated in the maintenance of mood, and that their depletion is *the* cause of depression. This relatively simple idea, now with many modifications, is still one of the bases of virtually all biological approaches to the treatment of depression.

Meanwhile, also in the 1950s, a Swiss company had developed a new drug called imipramine, which was intended as a treatment for schizophrenia, because it was chemically very similar to chlorpromazine (Thorazine), the first effective antipsychotic. It was ineffective against psychosis, but was mood-elevating, and was eventually tested successfully for depression. At this point two of the major classes of antidepressant had been discovered—Monoamine Oxidase Inhibitors (MAOIs) and Tricyclic Antidepressants (TCAs).[15] A great deal of research and development followed, and new MAOIs and TCAs were tested and marketed, and showed

---

[13] There is still a sizeable cohort of psychiatrists and other therapists who are not at all happy with these developments, and refer to this paradigm rather disparagingly as 'the medical model'. The rest of this chapter should show why this disparagement is ill-founded.

[14] The story of the discovery of these and other antidepressant drugs is well told in Healey 1998. See also the excellent chapter on antidepressants in Kramer 1994.

[15] The name comes from their structure: they are built on a scaffolding of three rings of carbon atoms.

considerable efficacy. The MAOIs were used rather less—as they still are—because of cardiovascular and other side-effects, and the TCAs and some related drugs became the mainstay treatment of depression until the 1980s, though they were often poorly tolerated and could be fatal in overdose.

While the MAOIs and TCAs were effective against depression, there was nothing available for mania except sedatives (e.g. barbiturates or antipsychotics), and nothing at all for controlling the cycling of bipolar disorder. The first effective antimanic and mood stabiliser was also discovered by accident. In 1949 the Australian psychiatrist John Cade was testing the hypothesis that high levels of blood urea caused mania. For this work, he needed a soluble salt of uric acid, and injected guinea-pigs with lithium urate. The results were precisely the opposite of what he had expected: the guinea-pigs became lethargic, not manic. He ran some clinical trials of lithium salts with manic patients and published his results, and in 1954 the Danish psychiatrist Mogens Schou undertook a series of successful trials.

It took nearly two decades before lithium was generally accepted as a clinical resource; it was not widely used until the 1970s, but has turned out to be a highly effective mood-stabiliser. How it works is not yet fully understood, but it does have detectable effects on neurotransmission. Cade's accidental discovery led to the development of the most widely used and effective antimanic drug, and the first of the quite small group of drugs that can control cycling. By the 1950s, then, both poles of the mood disorder spectrum, and the alternations between them, had shown themselves susceptible to purely chemical intervention.

So depressive illness can be relieved by medications with known effects on particular biochemical systems (however primitively understood). Therefore depression, whatever the sufferer's psychological experience, is at bottom a 'medical' condition like any other. It must be at least proximately a neurochemical illness, since dysregulations of particular neurotransmitter systems can be targeted for symptomatic relief. We have learned a great deal since the 1950s; the current picture is much more subtle, complex and confusing. It is not just the amount of neurotransmitter present that is involved, but the number and type of receptors, differential sensitivity of pre- and postsynaptic receptors, and complex interactions between various neurotransmitters and between them and hormonal systems and various proteins in the brain, and between all of these and the outside world; and there is a major (less proximate but crucial) causal role for the stress and immune responses.

But one basic fact appears to remain, no matter how much we learn. The only effective and relatively quick relief of the major symptoms of depression is achieved by treatments based on a theory which makes no reference to a mind existing independently of the physical brain.[16] Mind disease *is* brain disease—however unclear the relations between the two may be.

[16]This applies as well to electroconvulsive therapy (ECT), which is faster than medication. See below.

The biology of depression

> [...] "biology" is not some substance that is segregated or localized inside the initial state of the organism at birth, circumscribing the domain to which evolutionary analyses apply. It is also in the organization of the developmentally relevant world itself, when viewed from the perspective imposed by the evolved developmental mechanisms of the organism. Thus, nothing the organism interacts with in the world is nonbiological to it, and so for humans cultural forces are biological, social forces are biological, physical forces are biological, and so on. The social and cultural are not alternatives to the biological. They are aspects of evolved human biology and, hence, they are the kinds of things to which evolutionary analysis can properly be applied.
>
> —J. Tooby & L. Cosmides, The psychological foundations of culture (1994)

The late-onset paradox

For a little over half a century scientists studying the biology of mood disorders have largely agreed that dysregulation of amine systems is the central mechanism. In the beginning, the fact that drugs that increased amounts of free amines gave symptomatic relief was taken to illustrate the proximate cause: what we might call the 'deficit theory'. It looked as if things were rather simple after all: getting more amines into the synapse should cure depression.

But research and clinical experience kept bringing up complications. One of the worst was the delay between the onset of reuptake inhibition and the onset of antidepressant effect. The theory predicts that once there is (say) more serotonin in the synapse the symptoms should start vanishing. But a drug like Prozac, for instance, may produce an increase of serotonin in the synapse within a day or so after the first dose, but usually takes anywhere from two to six weeks to show any antidepressant effect. The side-effects however come on almost immediately, often as early as the first day. Such delay occurs with most antidepressants, for most patients. (If the desired mood changes show up within the first few days, as does happen, we might suspect placebo effect, not the drug working.)

Why should this be? I will tell the story in pieces in the rest of the chapter, as it is part of a very complex cascade of chemical reactions within the brain and the individual neuron. But we can dispose of the simplest (for a while thought to be the only) reason now. When a cell that requires a particular chemical to function properly is deprived of it, it often 'upregulates' its receptors for it. To be anthropomorphic, the cell cannot find enough of the desired substance, so in an effort to get more it makes more receptors. Of course this is counterproductive: if you do not have enough of what you need, making more receptors is going to make the situation worse, not better, because there will be less to go round. But this is the only thing a downstream neuron can do; neural signals propagate in one direction. If a neuron is starved of a neurotransmitter and makes more and more receptors, the best way to make it stop is to satisfy its craving; and this is apparently what most antidepressants eventually do. They keep more neurotransmitter in the synapse by disabling the reuptake pump (TCAs and most modern drugs) or stopping the degradation process (MAOIs); and eventually the postsynaptic cell decides that there is enough,

and begins to dismantle ('downregulate') the extra receptors. It takes a fair amount of time and energy to build or dismantle receptors; and until there is a match between the amount of neurotransmitter and the number of receptors, there will be no 'balance', and the symptoms will not start to remit.

This theory stood as a placeholder for quite some time; but now that more is being learned about the internal chemistry of neurons, it is no longer adequate as a sole account. One problem is that downregulation of receptors also often occurs some time before antidepressant effect, even simultaneously with neurotransmitter increase. In the next few sections I will explore some of the less well-known markers of depression, and then begin to construct a causal chain that will incorporate the more sophisticated chemistry and the so-far neglected stories of life-experiences and genetics: depression may seem to fall from the air, but there is usually a long history behind it.

Biological markers: anatomy and metabolism

For more than two decades the depressed brain has been subjected to imaging studies (MRI, PET, etc.). The results of over 12,000 studies (Mamo & Kapur 2008) have been that depression has significant brain-level correlates. Blood-flow and glucose metabolism, receptor-occupation by neurotransmitters, density of receptors, even the gross anatomy of the frontal lobes, hippocampus and amygdala of depressed patients, differ from what would be expected in the nondepressed. Here are the results of an early but clear and relatively simple study (Dreverts *et al.* 1997). This reports PET (positron emission tomography) studies of blood flow and glucose uptake, and CT scans of relevant portions of the prefrontal cortex in depressed (unipolar and bipolar) patients and normal controls.[17] This paper concentrates on regions of the prefrontal cortex which control interpretation of emotion and response to stressors, and are particularly sensitive to amine neurotransmitters. Visible lesions in these regions (tumours, strokes, trauma) typically produce 'flatness' or lack of emotional response, impaired concentration, lack of motivation, and other depression-like symptoms. The main findings were: (a) depressed patients showed lower blood-flow and glucose uptake in the relevant areas than non-depressed controls; and (b) there were marked anatomical differences as well, depressed patients displaying considerable reduction of grey matter (neuronal cell-bodies) in at least one hemisphere. The most characteristic volume reductions are in the hippocampus and frontal lobes. A meta-analysis (statistical summary study of other studies) shows average 8% volume reduction in the left hippocampus and 10% in the right; and the amount of reduction appears to correlate with the number of prior episodes (Videbech & Ravnkilde 2004). There is also evidence for volume reduction in the amygdala, which mediates fear, anxiety and similar emotions (Pezawas *et al.* 2005). So even at the level of

[17]Oxygen- and glucose uptake are indexes of brain activity; the more work a particular region is doing, the more fuel it needs, hence the more glucose and oxygen. For discussion of recent imaging studies see Mayberg 2006, Mamo & Kapur 2008.

physical structure and visible metabolic activity, depressed and non-depressed brains look different.

Investigations of blood-supply to the frontal lobes of depressed and non-depressed patients, and more convincingly, the same depressed patients before and after antidepressant treatment, reveal a similar picture. Depressed frontal lobes have less blood-flow, i.e. deficient oxygenation, and non-depressed or treated depressed frontal lobes show increased blood-flow. Depressed patients also show unusual activity in the limbic system.[18]

Other aspects of brain function provide further clues to what might be going on in both unipolar and bipolar depression. The most interesting perhaps is the association with the right/left specialisation of the cerebral hemispheres. For instance, affective disorder often follows damage to the frontal or temporal lobes from strokes and other trauma. As Goodwin & Jamison report (1990: 507), 'most frequently associated with depression or mania were lesions located in the frontal or temporal lobes. Left-sided lesions tended to be associated with depressions and right-sided lesions with symptoms suggestive of mania'. Other studies suggest a disruption of the normal functional relation between the right and left hemispheres in depressive illness. This data is not yet clearly understood, but it strongly suggests some kind of structural localisation in mood dysfunction.

Biochemistry: amine dysregulation

The brains and nervous systems of depressives (at least during depressive episodes) are chemically different from those of the non-depressed. For example, it has been known for a long time that suicides, on autopsy, tend to show below-normal concentrations of serotonin and its breakdown products in their cerebrospinal fluid, and abnormal distribution of certain receptor-types. And since the commonest cause of suicide is generally agreed to be depression, here is a potential (if weak) corroboration of the amine theory.[19] And nearly all effective antidepressants, whatever else they do, act either directly or indirectly on serotonin reuptake or degradation. There is another causal sequence lying behind the dysregulation of serotonin, involving the stress response; I will return to this in the next section. At this point I want to look at the two other neurotransmitters involved in depression, noradrenaline and dopamine.

The evidence for their role is somewhat less direct than that for serotonin. The picture is

[18] See Goodwin & Jamison 1990: ch. 18, and d'Haenen 1997. There is a stunning colour plate in Goodwin & Jamison (Fig. 18-1), showing PET scans of glucose uptake in a rapid-cycling manic depressive patient. For an imaging study showing ways in which the structural anatomy of the depressed brain predicts antidepressant response see Costafreda *et al.* 2009.

[19] This evidence, though frequently cited, may be ambiguous; see Smith & Cowen 1997. It is clearest when the suicide is violent, and associated with a violent, aggressive and impulsive personality. Smith & Cowen provide an excellent if highly technical coverage of what is known about the role of serotonin in depression. The best evidence for the correlation of serotonin concentrations with mood comes from baboons, which are highly social and easily studied animals, and show depressive behaviour in many of the same kinds of circumstances that humans do. For a thorough account and references see Sapolsky 2004.

further complicated by the fact that serotonin exerts a modulating influence on these chemicals as well, through projections from their sources in the more ancient parts of the brain. First, noradrenaline. As we saw in the last chapter, its primary functions involve arousal and alertness, and response to external or internal threat. There have been numerous observations of noradrenaline depletion in depressed patients and other primates. E.g. reserpine, which depletes presynaptic noradrenaline, precipitates depressive symptoms in nondepressed controls, and many antidepressants reverse the effects of reserpine. Contrariwise, some antidepressants that inhibit noradrenaline reuptake (even if they have no effect on serotonin) are effective against certain types of depression. It has also been shown that one of the most consistent effects of long-term antidepressant treatment is reduction in the number of postsynaptic noradrenaline receptors, suggesting that the therapeutic effect involves reversing an original shortage of noradrenaline. .

Dopamine, as we have seen, is involved in the limbic 'reward' system; hypothetically a decrease could lead to loss of pleasure and motivation (the typical depressive anhedonia). On the other hand, an excess could lead to excitability and increase in goal-seeking behaviours: thus dopamine would be a good candidate for involvement in bipolar disorder. There is some biochemical evidence to support this as well: reduced production or excessive metabolism of dopamine is involved in depression, especially when it is marked by psychomotor retardation—but not in agitated states or manias. (In fact some dopamine agonists used in treating Parkinsonism may produce mania as a side-effect.) Depletion of dopamine is associated with lowered mood, and excess with elevated mood. One of the most activating of all antidepressants, Wellbutrin, is a noradrenaline and dopamine reuptake inhibitor which does not affect serotonin, and a major sign of its dopaminergic activity is the fact that it can be used to help restore sexual function damaged by other antidepressants or by depression itself.

Stress and amine dysregulation

It now seems clear that stress is one of the major causes of amine dysregulation.[20] The causal pathway is somewhat indirect, and works primarily through the stress-reaction chain, or HPA. In summary, severe stress makes the hypothalamus release hormones that signal the adrenals to produce first adrenaline and then cortisol.[21] Cortisol, in addition to mediating the long-term stress response, can, with excessive exposure, have complex effects on the brain. Among other things it dysregulates serotonin and other amine pathways. The shortage of amines in the synapse is largely a secondary effect of stress, due to long-term high cortisol levels in the brain. Excessive cortisol can damage neurons, especially in the hippocampus. And at least one of the genes that is involved in the serotonin system has sites that bind cortisol, which establishes a connection at

[20]For a summary of the evidence see Sapolsky 2004, especially chapter 14.

[21]Actually it produces a number of chemically related hormones called glucocorticoids, but cortisol is the most important and I will use it to stand for the rest. For a thorough account of the role of glucocorticoids in stress reactions, Sapolsky 2004.

the micro-level (Pezawas *et al.* 2005). I will return to the significance of this gene later. There is also evidence that a neuroprotective protein called Brain-derived Neurotrophic Factor (BDNF), which is necessary for maintenance, growth and regeneration of neurons, is inhibited by cortisol. This inhibition may be the final step in the chemical cascade leading to depression; it is also significant that it is reversed by many antidepressants.[22]

So given the interactions discussed above, one cause of neurotransmitter dysregulation appears to be stress, whether exogenous or endogenous. Recent research is finding increasing evidence for this. Both the adrenals and pituitaries of depressed patients tend to be enlarged, and there are abnormally high levels of CRH (the hormone that causes the adrenals to secrete cortisol, and also acts on the amine systems) in their cerebrospinal fluid. On autopsy, they tend to show more CRH-producing neurons than normal in the hypothalamus, and signs of stress-induced damage to the hippocampus.

This further supports what has for some time been the standard account of the genesis of mood disorders— the so-called 'stress-diathesis' model ('diathesis' = predisposition). Depression and related disorders result from the effects of chronic stress on neurotransmission and endocrine systems, superimposed on innate (genetic) predisposition or vulnerability. Amine dysregulation leading to BDNF suppression is the most proximate cause we know, and stress and its consequences the next one back in the chain of causation.[23]

Recently more light has been cast on the relation between stress and depression, via the mediation of the HPA and the immune system. The epidemiological relation is fairly clear: it has been shown time and again that childhood abuse, poor parenting, neglect, and other early (or even later) trauma are risk factors for the development of mood disorders. One recent study has found that women with a history of childhood abuse have a four-fold greater risk for major depression than women without this history.[24]

Finer scrutiny of the amine and CRH pathways shows an even more intimate connection with stress than previously supposed. CRH neurons in the hypothalamus project to the source of noradrenaline in the brainstem, so that stress can lead to hyperactivation of the noradrenaline

---

[22]Shimizu *et al.* 2003. The role of BDNF has been shown strikingly in its ability to stimulate the regeneration of cells after brain damage (Wilson *et al.* 2008). The effect of antidepressant treatment on adult hippocampal cells (in a way implying BDNF stimulation) has been shown in Wang 2008. This is a mouse study, but regardless of the ethics mice have proved a useful model for human depression.

[23]For a clear and well referenced account of the chemistry of depression, including a good treatment of BDNF, see http://www.psycheducation.org/mechanism/MechanismIntro.htm.For the stress connection in particular the most recent survey I know is Bartolomucci & Leopardi 2009.

[24] See Kendler *et al*. 1993, McCauley *et al*. 1997, Mullen *et al*. 1996. However complex this whole account seems, it has really been just a sketch of a few of the many mechanisms probably involved in the genesis of different types of depressions, or even all depressions. For a detailed summary of our current knowledge of the structural, chemical and genetic factors involved in depression, see Krishnan & Nestle 2008.

system. This is probably why anxiety so typically co-occurs with depression, and why many antidepressants also have anti-anxiety effects. Both the CRH and noradrenaline systems are known to become hyperactivated after early-life trauma in humans and other primates.[25] Depression is often associated with immune dysfunction. Depressed patients tend to be more susceptible to infection than the nondepressed, and stress in general predisposes to raised cortisol and diminished immune response, and hence to infection.

Recent work on the immune system and depression reveals a complex, bidirectional relation. Stress enhances some immune functions and reduces others, and the activated immune response, through the release of certain 'messenger' chemicals called cytokines, can produce symptoms of stress and depression. The micro-level picture of depressive biochemistry looks rather like a combination of inflammatory disease (immune hyperactivation) and immunosuppression.[26]

Depression and sleep architecture

Another clear instance of the brain- rather than mind-centredness of depression is its effect on sleep. There is more to the sleep disorders accompanying depression than just insomnia brought on by the nasty contents of the mind; depressive sleep has a characteristic physiological signature. Sleep is not just shapeless unconsciousness; it has an 'architecture', a sequence of well-known states. A night of typical sleep is divided into two basic types: one characterised by rapid eye movement (REM sleep), and another by lack of such movement (NREM sleep). An 'average' night's sleep appears to be constructed this way:

(a) NREM Sleep

STAGE 1. 'Drifting off', dream-like imagery, normal vital signs (heart-rate, blood-pressure, respiration, temperature). Sleeper easily aroused.

STAGE 2. Irregularities ('sleep spindles',short irregular bursts of electrical activity) enter the EEG pattern. Arousal is more difficult than in stage 1.

STAGE 3. Transition to deep sleep. Vitals drop, skeletal muscles relax. Some dreaming may occur in this stage, which on average is reached about 20 minutes after stage-1onset.

STAGE 4. Usually referred to as slow-wave sleep (SWS). The 40Hz neuronal 'background

[25] See Maes 1997, which shows that depression shares many biochemical properties with 'acute phase' immunological response, and has many of the features of an inflammatory disease, along with immunosuppression. For more recent work, Sadek & Nemeroff 2000.

[26] This account is based on Sadek & Nemeroff 2000.

oscillation' stops. Vital signs drop to normal minimum, though gut motility increases. Skeletal muscles are relaxed, but motion is still possible, and sleepers tend to change position quite frequently. Arousal is difficult, and an aroused stage-4 sleeper is usually confused. This is the stage in which sleepwalking is most likely to occur.

(b) REM Sleep.

This normally begins about 90 minutes after stage 1 NREM onset. There are sudden EEG changes, and the sleeper appears to 'backtrack' through phases 4-1. Temperature and heart-rate rise, gut motility decreases, and oxygen consumption is greater than when awake. The eyes move rapidly under the closed eyelids (hence the name), but there is temporary paralysis of skeletal muscles. Most dreaming takes place during this phase. Typically NREM/REM cycles occur about every 90 minutes; after each REM episode there is a return to stage 4 NREM sleep. REM periods usually lengthen as the night progresses, starting out at roughly 5 minutes and ending up shortly before waking at nearly an hour.

Neurotransmission and hormone-release patterns alter during the different sleep phases. During deep sleep, particularly SWS, there is a rise in serotonin levels and a decrease in noradrenaline; in REM sleep, there is an increase in noradrenaline.

In depression, this architecture is disrupted in a characteristic way. First, the onset of sleep itself is delayed, and overall sleep-time is decreased, as is efficiency (there are frequent intrusions of wakefulness into deeper sleep). More specifically, the expected ratios between REM and SWS are virtually inverted. There is decreased 'REM latency' (faster onset), and more and 'denser' REM sleep than normal (faster than usual eye-movement). There is also an unusual asymmetry between EEG readings from the right and left hemispheres. These signs are partly corrected in patients undergoing antidepressant treatment. Other abnormalities have been recorded: a rise in temperature, increased ACTH and cortisol output, and decreased output of hormones expected to be at high levels (e.g. thyroid-stimulating hormone, testosterone, and growth hormone). While other psychiatric illnesses (e.g. dementias) also show sleep changes, they are not the same kind, and the particular pattern described here may even be diagnostic for major depression.

Depressed patients, particularly those who also suffer from anxiety, are likely to have sleep disorders (current estimates are that 85% have insomnia). Even depressives undergoing treatment are rarely able to sleep without some kind of help (unless they are taking sedating antidepressants). But most hypnotics, while they do get you to sleep, also disrupt sleep architecture. Alcohol and barbiturates suppress REM sleep, but not SWS; benzodiazepines tend to reduce SWS. In either case sleep is likely to be less restorative than if it was not at least partially drug-induced; but it is better to be able to sleep than not.[27]

[27] On sleep and depression see the detailed treatment in Goodwin & Jamison 1990: chapter 19, and Berger & Riemann 1993.

Biological rhythms and cycling

> Surely among the most abstract of concepts, the rhythmic organization of time is grounded in material reality. Biological functioning is itself organized into periods linked to the rotations of the earth about the sun and the moon around the earth [...] Although biological rhythms cycle in synchrony with these "celestial mechanics," they are clearly regulated by endogenous processes.
>
> —F.K. Goodwin & K.R. Jamison, *Manic-depressive illness* (1990)

The architecture of sleep is rhythmic—i.e. particular stages repeat in a predefined sequence, and the decreased REM latency and REM/NREM reversal characteristic of depression is clearly a disturbance of rhythm. The colloquial meaning of 'rhythm' is intuitively obvious, but many biological phenomena are rhythmic in a rather special sense. Let us define rhythm as 'relatively regular recurrence'. A random list of biological phenomena that could be called rhythmic would include: the beating of the heart; mood-cycles in bipolar or recurrent unipolar disorder; the menstrual cycle or the recurrence of oestrus in other animals; hormonal changes during sleep; the timing of germination, flowering and seed-setting in plants; the timing of insect and amphibian metamorphosis; the onset of hibernation in bears; changes over the day in blood-sugar levels; rises and falls in libido; the 11-year cycle of predator and prey populations (possibly connected with sunspot activity).

Some periods are regular: menstruation in humans runs on a lunar clock, while hibernation and flowering run on solar clocks (the period is the year). But there are irregular periodicities as well, particularly in mood-cycles (as well as epilepsy and migraine); these may not be keyed to 'external' clocks, but to more complex and poorly understood internal period-setters.

The most important clocks are those that run on a roughly 24-hour cycle, the circadian ('about a day') rhythms: a typical example is the sleep-wake cycle, which when unimpeded has a period of around 25 hours, though in modern industrial societies we tend to shorten it by the way we organise our days. There are also infradian ('below a day') clocks, like those controlling blood-sugar and hormone levels; and ultradian ('beyond a day') clocks, like the lunar and solar cycles.

Some characteristic rhythms and periodicities in mood disorder can be illustrated by my own depression. These are difficult to observe when you're in the middle of them, but after years of experience and a good deal of deliberate attention, I think I have a fairly clear picture, and feel confident about the following description. When I am in remission, the only rhythmic disturbances apparent are in sleep: even with medication I usually have trouble falling asleep no matter how tired I am, I start dreaming early, and wake early. (This is partly due to disruption of the normal rhythm by alcohol.) But when the disease breaks through, and a new cycle begins, different patterns appear. One of the first, very often but not always, is advancement of the circadian clock: my 'morning' begins about an hour earlier each day for a period of three to five

days. I wake up around 7.00, then the next day at 6.00, then at 5.00, then at 4.00.[28] When I start waking regularly at 4.00 and cannot get back to sleep, I know that a new episode is on the way (how bad or how long or what kind I cannot predict). A day or so before this retiming begins, I often have a feeling of vague disquiet, a kind of 'aura', a sense of foreboding and edginess, sometimes vague nausea and tremor. So the first hard sign that a new mood-cycle is beginning is often temporal: my internal sleep-wake clock becomes less and less 'entrained' to the circadian light-dark cycle, more autonomous, and eventually begins a 'free run': the clock becomes self-regulating, and disregards the external factors like temperature and light that would normally cue it.

Once I go into a full depressive episode, the overall mood landscape re-entrains itself to a new circadian rhythm: I wake early, in an affectively neutral state, but within about fifteen minutes I become more and more depressed and anxious. This persists for much of the morning, on some days into the afternoon, at which point my mood begins to lift, and may go back to 'normal', or only slightly dysthymic, or even hypomanic. Then there is another drop in mood in the afternoon, which may remain until evening, or sometimes persist until (and in dreams after) I go to sleep. When this kind of episode starts, it tends to go in whole-day units: I have largely depressed days, or sometimes the depression is replaced by anxiety and panic, or by black hypomania, or even by euphoria. That is, the episodes that are signalled by sleep disturbance may be simple recurrences of depression, or the beginnings of rapid-cycling bipolar sequences, or even long-lasting hypomanic states —but depression is almost always the first stage.

It is difficult to predict when these episodes will occur; but they are partly linked to an ultradian clock, since there are times of the year when they tend to be worse. The ratio of light to darkness (the 'photoperiod') certainly plays a part, as does the day-to-day rate at which the photoperiod changes (the closer to an equinox, the faster). I am at my most unstable in summer and early spring, and most likely to have long and suicidal depressions in the winter and late summer, and euphoric hypomanias in late spring and early summer. Occasionally the mood-shifts reverse, and I get manic at the onset of winter.

This effect is particularly clear when I travel, since most of my travelling is not of the jet-lag east-to-west type, but across the equator. So for instance if I fly from South Africa in May or June (our winter) in a rather depressed state, and arrive in European summer, within a day or so the clock seems to reset to the new photoperiod, and I tend to be normal or hypomanic most of the time. The reverse often happens when I fly south again: this is partly what lay behind my wife's comment when I phoned her from Cambridge that when I got back to South Africa the shit would hit the fan again.

[28]This advancement was much commoner when I was younger; as I age the cycling tends to become more disorderly and less obviously controlled by a circadian clock.

Proximate causation: reprise

The material in this chapter so far converges on a single critical point: whatever may have started a depression in the first place, the disease itself presents with the biochemical signatures of stress-related illness, immune dysfunction and amine and limbic system dysregulation. And there are physical signs in the brain and bloodstream, detectable by the standard procedures for investigating human biochemistry, brain function and structure, as well as characteristic sleep anomalies. This implies that since some signs of depression are detectable by non-'psychological' means, some treatments should succeed by targeting the anomalies that produce these signs. Whatever the higher-level (cortical or experiential) correlates of depression may be, stress-induced amine and HPA dysregulation appears to be involved in the final common pathway. All the endocrine dysregulations and anomalies in brain function (glucose uptake, lateralisation, cell loss) seem to be mediated by these systems. Disordered neurotransmission is the business end of mood disorder.

Life stories as causes

> But indeed nothing happens at one moment rather than another. The history books will make it much more definite than it is.
>
> —Virginia Woolf, *Diaries*, 18 February 1921

Escaping dichotomy

What is experienced psychologically as depression manifests physically as a distinct cluster of signs. Our understanding is grossly incomplete, but at least we have some information. A level of proximate causation seems to have been identified, which suggests where our primary therapeutic energies should be directed. But the stress connection, the epidemiology and the narratives that emerge from talk therapy suggest that we can go back a step. Life-stories must play a part in many if not most chronic depressions, since it is here that we find so many of the predisposing stressors. Can we tie together the 'mental', 'environmental' and 'historical' data and theory with what we know of the neurophysiology? We must eventually: it would be outlandish if the affective and cognitive symptoms and often the known histories of stress were unconnected with the biological markers. How can a psychological insult be transformed into or manifest as a physiological dysregulation, which in turn manifests as a psychological disease? If the biographically ultimate cause of a particular case of depression may be some kind of life event, e.g. bereavement, abuse, dysfunctional relation with a parent, there ought to be a coherent causal story.

But often this necessary theoretical coherence is ignored, or sidestepped by a dualist either/or exclusivity. Here is a characteristic example, from a popular book by the clinical psychologist Lauren Slater (1997: 65). The topic is schizophrenia, but the same misjudgement

occurs commonly in discourse about depression.

> No one knows for sure why the schizophrenic has such a hard time with words, why so little of his language makes sense. Ask him how the weather is and he might tell you, *Frogs be flying a green way* [...] Is this mumbo jumbo due to a dysfunction of the parietal lobe [...] or to some other kind of neuronal collapse? Or is it due to the schizophrenic's mother, who, early on, tongue-tied him with an overbearing love? People cannot definitively say.[29]

This reads almost like a parody of the extreme positions in the therapeutic and theoretical 'causation wars'. The psychoanalyst or psychodynamic psychologist targets (what may be) the ultimate biographical cause, and hopes to reduce the symptoms by disabling the original stressor through recall, understanding, cultivation of insight, etc. The psychopharmacologist tackles the symptoms themselves, regardless of what may have caused them, often without interest in possible causes, or even in the face of the impossibility knowing anything reliable about them.

I would argue that in many if not all cases, certainly with respect to depression itself (not its effects on the rest of living), the ultimate biographical cause is usually therapeutically irrelevant or nearly so.[30] It does not matter a whole lot if you broke your arm because you were inattentive and tripped on the stairs, or if a mugger broke it. A broken arm is a broken arm, and the emergency doctor will treat it as such, rather than letting you lie there and suffer while he asks you what deep-seated family conflict or sight of something nasty in the woodshed made you miss a step or wander into the wrong neighbourhood.[31] The question of which kind of repair is best for a particular broken neurotransmitter system may still be open, but good psychiatrists will judge with care which one (often both) might be appropriate in a given case.

---

[29]Incidentally, it is rather odd for a professional psychologist to attribute language problems to the parietal lobe, which is not involved. Perhaps she meant 'temporal' or 'frontal'?

[30] The major exception is where the ultimate cause is constantly present, e.g. a continuing abusive relationship or otherwise intolerable home life, loneliness, unemployment or extreme poverty. In such cases no treatment may be really effective until the environment is changed, since every therapeutic improvement will be undone by a return to the causal context.

[31] Fieve (*Moodswing*, p. 201) notes that many patients are not interested in 'insight' into what caused their depression, but simply want symptomatic relief. He suggests that no more than about 30% of his patients are interested in a psychoanalytic or other type of understanding. These are the ones he refers to psychotherapists *after* they are stabilized on drugs. In a quantitative study of patients undergoing psychotherapy in Maryland (reported in *Psycholink* 3:10, November 1994, p. 8), 'few patients said they benefited from the psychoanalytic insights of the therapy'. The conclusion was that 'most patients undergoing psychotherapy do not want the type of lengthy, psychoanalytic, insight oriented psychotherapy favored by therapists'. Stuart Sutherland , himself a trained academic psychologist, notes in a number of places that the 'insights' gained through psychotherapy did him no good at all, though the ones he believed were of mild theoretical interest. I return to th is matter in detail in chapter 5.

Life-events, stress and kindling

We have enough information now to trace, at least tentatively, the pathway from stressor to disease without dichotomising. Given our current state of knowledge and ability to intervene, the proximate cause of chronic depression is a disorder in one or more neurotransmitter systems, which affects brain metabolism and function. Your childhood (or other) experiences may, through imposing severe stress, have dysregulated your brain. It is your 'psychological' brain's reading of its 'physical' counterpart's malfunctioning chemistry and/or disordered structure that produces the effects or experiences we call depressive illness. In terms of brain function, the causal stressor, perceived by both the limbic system and higher cortex, has been transformed into a subcortical dysregulation that sends abnormal messages both to frontal cortex and to the rest of the body; these are what is experienced as depression.

We need then to articulate an inclusive middle way between purely psychodynamic or cognitive and purely neurochemical approaches; this ideally would allow the full richness and humanness of mood disorder to surface, and still be good reductionist, nondualist science. We need to know how, if depression can ultimately be caused by an early bereavement, a dysfunctional family, a pervasive feeling of lack of control or a monstrous childhood, this can become a chronic illness recurring *in the absence of the original stressor*; and how it can be reflected in and successfully treated as a neurochemical dysregulation, without reference to its history.[32]

One influential view of how this comes about is Robert Post's 'kindling' theory.[33] This is a well-argued but still somewhat speculative idea, since it has not been empirically demonstrated; but the argument is so good that many investigators accept it at least as an interim account. As the name suggests, kindling is an incremental process. The idea first occurred in some animal experiments involving epilepsy induced by the electrical stimulation of particular brain regions. One kind of seizure is essentially due to a 'weak point' in the brain's circuitry that responds to some stimulus with neuronal hyperexcitation and continuous, uncontrolled firing. As the stimulus is repeated over time, the relevant brain region becomes increasingly sensitised to it. The seizure-threshold is lowered, so that it takes progressively smaller stimuli to provoke a seizure. Finally, this region becomes so hyperreactive that its circuitry locks onto some periodic mechanism, and the seizures become autonomous. The result is a full-blown recurring epilepsy that goes its own way without the need for any stimulus.

Arguing from an elaborate and sophisticated analogy, Post suggested that a similar

---

[32] The following should be read with one caveat in mind: the causal chain includes a pre-existing vulnerability to depressive reactions to stressors. Not everybody going through the same experiences reacts the same way, and what eventually throws one person into a suicidal depression might be shrugged off by another. The hidden variables are temperament and other genetic predispositions (see below).

[33] Post 1992. There is a very detailed and lucid discussion of kindling in Kramer 1994: chapter 5. For critical discussion and the suggestion of other, related mechanisms, see Monroe & Harkness 2005.

kindling may produce chronic depressions.[34] Say you are constantly subjected to unavoidable and irremediable stress by events in your childhood that lower your self-image and sense of status and worthiness, and put you in a constant state of (perceived) 'danger'. An example would be abuse, physical or psychological or both, by a tyrannical parent or group of peers. We know from work on other social primates, both in the wild and in captivity, that status in their very hierarchical societies is closely related to free serotonin concentrations (Sapolsky 2004). That is, dominant animals have higher serotonin levels than ones lower in the hierarchy, and a drop in status (say being deposed from top baboon to somewhere down the scale) is associated with lowered serotonin and raised cortisol levels. And giving antidepressants to low-status troop-members can cause them to become dominant. As a kind of Just-So story ('How the depressive got his amine problems') let us project this information to the dominance (and therefore status) relations in a human family. (This is not Post's argument but my own, based on his theoretical framework.)

Consider a child—by definition low in the status hierarchy of his troop, to retain the baboon image—who is constantly attacked by his father, either physically or verbally. (Not that baboons attack verbally, but threat-displays as opposed to actual attacks play a similar role.) Nothing the child does is right, every time he achieves something it is not good enough, nothing he does seems to elicit approval from the top baboon, and what appear to be 'good' acts provoke apparently irrational rage. The setting for my parable is autobiographical: my childhood and adolescence were largely consumed by wrestling for status and self-image with an alpha baboon, who happened to be my father. In addition, and this is not irrelevant to the primate-troop, I was unfortunate enough to be short, smart, middle-class, bespectacled and unathletic in a primary school with a socially mixed intake. In such a setting, this collection of attributes put me on the wrong side of the class war: I had to spend a lot of time and energy dreaming up alternative, often bizarre and complicated routes home from school, to avoid the knots of apprentice thugs I would otherwise meet, and so maybe pass another day without harassment or physical damage.

Now each time a child with the requisite vulnerability is on the receiving end of the appropriate negative stimulus, he reacts psychologically; and the physical reflex of this reaction (or the reaction itself) is stress, leading to a prolonged 'fight or flight' reaction, and attendant CRH- and cortisol-mediated dysregulation of the serotonin system and sensitisation of noradrenaline pathways. There is less serotonin production, faster reuptake, diminution of postsynaptic receptor sensitivity, growth of excessive numbers of receptors. This will almost certainly be accompanied by changes in neuronal wiring, including death of hippocampal neurons, as well as alterations in the internal gene activity of the neurons themselves. Repetition or continuation of these insults will make the system increasingly sensitive; smaller and smaller stimuli will set off what is physically reflected as lower serotonin levels (with all the effects this may have on other amines and BDNF) and psychologically as feelings of unworthiness, guilt, depression and anxiety. Eventually the system becomes so exquisitely responsive that it no longer

[34] There is further evidence for this in the efficacy of anti-kindling drugs normally used for the treatment of epilepsy in certain forms of bipolar disorder.

requires any stress input—each episode kindles the next until the relevant circuitry locks into a permanent depressed state, or becomes hyperreactive. Then either very tiny (and apparently 'ir-relevant') stimuli will set it off, or it will develop its own independent periodicity. In that case the victim will perceive depressive (or manic) episodes as untriggered, unrelated to anything except the brain's own peculiar logic. They 'come out of the blue', or 'fall from the sky', because the period-setting mechanism is not available to introspection.[35]

In fact, as Peter Kramer has noted (1994: 113), kindling 'appears to be a kind of learning, but a learning that can occur independent of cognition'. That is, unconscious limbic subsystems ('zombies' again) can learn habits without the rest of the brain having any particular investment in the process; the end-results are 'discovered' by consciousness after the learning is complete. One might characterise kindled depression as a kind of procedural memory ('knowing how' rather than 'knowing that'); like a musician, one('s brain) 'practices' being depressed, and gets better and better at it, until the behaviour becomes pure routine.[36]

It is uncertain whether uncovering the original insult (assuming it has been repressed), or trying to get the patient to deal with it and obtain some kind of narrative insight into its correlation with his current depressed state is useful. But what is clear is that the symptoms themselves can be treated independently of their history, and indeed ethically speaking must be. Defibrillate first, then find out why the heart stopped.

This interconvertibility between the (unperceived) neurochemical and the (perceived) psychological frees us from the 'mentalist' vs. 'physicalist' dichotomy. The 'mental' is the way the brain interprets its own structural and chemical landscape at the experiential level. The neurochemical and the psychological are the same thing from different points of view; we just talk about them for clarity in two languages, 'biologese' and 'psychologese'.

We can now rough out an integrated if preliminary model of the causation of depressive illness: at least some social, psychological and biochemical partial causes have been exposed. But there are important dimensions still unexplored: in particular the complex relation between 'temperament' and depression, and the possible evolutionary origins and known genetic factors involved. This may point the way to a more unified, if still woefully imperfect and incomplete picture.

---

[35] This is of course grossly oversimplified, both from the biochemical and psychosocial points of view; it is merely exemplary. For some detailed discussion of this kind of nexus, see van Praag 1997 and Sapolsky 2004.

[36] There is evidence for this view: apparently the notion of kindling is well supported, but after a certain number of kindled episodes the effect dies away and the episodes recur but do not become worse. See Kendler *et al*. 2000.

Genetic susceptibility: a pre-experience cause

Heritage

> They fuck you up, your mum and dad.
> They may not mean to, but they do.
> They fill you with the faults they had
> And add some extra, just for you.
>
> —Philip Larkin, 'This be the verse'

We have already seen how they can do this by the way they treat you. But they can also do it inadvertently, with the best will in the world, even if they are as close to perfect parents as possible. Not by their actions, but by what they are and where they came from. Depression, like schizophrenia, autism, personality disorders and many other psychiatric diseases, can be as much a prenatal heritage as the result of events in your own lifetime. To show this requires an excursion into genetics.

Inside the nucleus of each of our 10-trillion cells (except red blood cells, which have none) is a set of double-stranded helices of DNA. This contains all the information necessary to build us from a fertilised ovum, maintain us during our lives and allow us to interact with our environment. This is the human genome, our heritage from our parents. In all cells except sperms and eggs there are two nonidentical copies of the genome: a maternal and a paternal. One's genome is a family heritage—but it is an extended family. We get our genes from our parents, who got theirs from their parents, and so on back to the earliest *Homo sapiens*, to *Homo erectus*, to the first mammal ... and eventually back to the (unknown) common ancestor of all living things. We carry in every body cell a partial record of the history of life and its contingencies, even bits of genetic material not originally ours at all, but spliced into our genomes by viruses and other parasites.

In the popular conception, the genes' task stops when they have made us. They give us our shapes, organs, sexes, eye-colour, family resemblances; sometimes they give us diseases.[37] But most of our 30,000-odd genes are not just recipes for making us; they are constantly at work in the present. Genes are perpetually turning on and off, making proteins that construct tissues or chemicals, telling cells when to divide or when not to, giving them orders to self-destruct, organising 'housekeeping' tasks like getting rid of waste and releasing energy, running the immune system, timing our circadian clocks, making hormones and neurotransmitters and receptors—to give just a sample. The results of all these complex and minuscule chemical operations are larger-scale processes like development of a mature individual from the single cell of a fertilised egg, and maintenance of that individual during its lifetime. Genes are messages

[37] Actually there are no genes whose sole function is to give us diseases; it is particular *variants* of certain genes that can cause pathologies.

from the past, since you get them from your parents; but most of them are also part of the working present.

Excursus: Mendelian and non-Mendelian inheritance

I assume that readers of this book have a basic idea of 'standard' genetics: that for instance some genes are 'dominant' and others 'recessive'. Your complement of genes is your 'genotype'; the physical, biochemical and behavioural traits they produce in interaction with your environment are your 'phenotype'. Here is a simplified reminder. It is broadly true that with respect to eye-colour, brown is dominant and blue is recessive. So we might conceive an 'eye-colour gene' with two variants or 'alleles', one dominant *B*, and one recessive *b*. Assuming one gets a single allele from each parent, the possibilities are:[38]

GENOTYPE	PHENOTYPE
BB	brown
Bb	brown
bb	blue

Given a dominant/recessive pair, the dominant will normally 'cancel out' the recessive, so that *BB* and *Bb* produce the same phenotype. Someone with the genotype *BB* ('homozygous' for *B*) will have brown eyes, as will the 'heterozygous' *Bb*.[39] But *Bb* is capable of passing down both blue and brown eyes, depending on the other parent. Say a *BB* mates with a *Bb*. Each gamete (sex-cell) of the first parent will carry *B*; each gamete of the second parent will have *B* or *b* (sex cells contain only half the species' chromosomal complement). Since the meetings of sperms and eggs are random, the offspring of these parents could be *BB or Bb*; on average 50% will carry the recessive allele, but all will display the dominant phenotype. But if both parents are heterozygous *Bb*, a quarter of the offspring will be homozygous *bb*, a quarter will be homozygous *BB*, and half will be heterozygous *Bb*; but three-quarters of the offspring will have the brown-eye phenotype, and only one quarter the blue. This is what accounts for traits 'skipping a generation': heterozygotes are 'carriers' who do not show the recessive phenotype, but it can surface in a mating with either another heterozygote or a homozygote.

This is classical Mendelian inheritance, the kind we learn about in school biology courses; but it is not the only kind, or even the commonest. Genes can also behave in much more complex and less predictable ways:

---

[38]This is an expository oversimplification, but it is true in principle. In actual fact, blue is not always recessive to brown; there is continuous variation along a scale from pale blue to brown, connected with the amounts of the pigment melanin that gene expression produces. So two dark-eyed parents under certain circumstances can produce a blue-eyed child.

[39] A zygote is a cell derived from the fusion of a sperm and an egg (Greek 'yoked'), which by division becomes an embryo.

(i) *Polygenic Inheritance* or *epistasis* (Greek 'standing-upon'). A number of genes act in concert to produce a given phenotypic effect. The construction of one type of human dopamine receptor involves at least four genes on different chromosomes, each of which builds one of the four proteins that make up the exterior or 'docking site', and each of these may have a number of alleles with different effects. Probably most important traits are controlled by a number of genes, operating in concert.

(ii) *Pleiotropy* (Greek 'many turnings'). Single genes may have multiple, often apparently unrelated, phenotypic effects. A classic case is the recessive metabolic disease Phenylketonuria (PKU). Homozygotes for the recessive allele deposit a dark-coloured pigment in joints and elsewhere, leading to early arthritis; this variant also yields nerve damage, mental retardation, fair hair and blue eyes. Similarly a gene that codes for part of a receptor for some neurotransmitter can have phenotypic effects on every activity of that neurotransmitter, including its interactions with others. It is likely that most genes are pleiotropic to some degree.

(iii) *Genetic heterogeneity*. In many cases, similar or identical phenotypes can be produced by alleles of different genes. There are apparently at least forty different genes which have alleles producing hereditary deafness (Barondes 1998: 205). And many genes are not direct 'causes' of diseases, but simply induce varying degrees of susceptibility to them, partly mediated by environment and other genes. Juvenile onset diabetes—and almost certainly mood disorders—come under this heading.

(iv) *Incomplete Penetrance*. Penetrance is a population-wide measure of the degree to which a given gene is expressed. Possession of a particular allele does not always mean that the possessor will show its effects. For instance, there are many genetic disorders in which identical twins (who by definition have nearly the same genotypes)[40] do not both show the results of a given gene or gene-complex. If one of a pair of identical twins suffers from bipolar disorder, the other has a 79% chance of also having it. Now this is clearly a genetic effect, since the bipolar disorder rate in the population at large is about 1%; but the genetic underpinning in this case is only 79% penetrant—as opposed to Huntington's disease (see below), which is 100% penetrant: every individual carrying the allele for it gets it.

[40]There are two types of twins: monozygotic or 'identical'(two fetuses from the cleavage of a single fertilised egg), and dizygotic or 'fraternal' (separate fertilisations occupying the same womb at the same time). Identical twins ought to have more or less the same genomes; fraternal twins, regardless of the accident of timing, are no more genetically similar than any arbitrary pair of siblings. They share on average 50% of the parental genes, but not the same 50%. The standard view that identical twins have absolutely identical genomes is not entirely true (Bruder *et al.* 2008). This paper shows that identical twins can have different numbers of copies of the same gene. There are also other differences: they do not have identical fingerprints.

(v) *Variable expression*. Some genes (or mutations or gene-combinations) are expressed to varying degrees in different carriers. This is the case for skin-colour, and for many genetic diseases, which can appear in different degrees of seriousness.

Polymorphisms

Genes were first imagined as unitary objects arranged on the chromosomes like beads on a string. Since the unravelling of the structure of DNA in the 1950s, the picture has changed radically; genes turn out to be complex structures with previously unsuspected properties.

DNA consists of a 'backbone' of the sugar deoxyribose with phosphate groups attached, and a series of nucleotides or bases that carry the genetic code: there are some 6 billion of these in the human genome. The ability of the famous DNA double helix to replicate, and thus serve as genetic material, is due to the fact that each of the four bases or nucleotides, adenine, cytosine, guanine and thymine (A, C, G, T) can pair only with one complementary base: A pairs with T and G with C. This means that if a strand of DNA separates into its two component sub-strands, each strand can form a template for its own recreation (the bases are built out of precursors floating in the environment).

The 'meaning' of a string of bases, i.e. the genetic code, is based on nucleotide triplets: each of the 20 amino acids out of which proteins are made is coded for by one or more triplets. So for instance AAA codes for lysine, AAC for asparagine and CAG for glutamine. A gene, then, is loosely speaking a string of nucleotide bases, which can replicate itself on cell-division; functionally, its job is the manufacture of proteins by a complicated process involving copying itself and the sending of 'messengers' to other parts of the cell, which direct synthesis.

So a gene is a string of bases, at a particular location on a particular chromosome. But not all of the material in the sequence codes for protein. A gene itself consists of coding pieces called 'exons', and noncoding ones called 'introns'; the latter are 'edited out' by RNA[41] before amino acids are manufactured. There are also stretches of DNA which do not code but control the expression of the gene, such as 'promoter' regions, which start the process of expression (see below).

These complex structures are not fully stable; in the process of replication bases can be lost or replaced by others; or parts of exons (usually triplets for particular amino acids called 'codons') can over-replicate, and produce multiple copies within a gene. These are 'mutations'. A mutation can be neutral in effect; it can also be harmful (even lethal), or—much more rarely—beneficial. Natural selection is built on these differential effects. Any mutation is likely to produce a different protein from the original gene; and some of these variant proteins can have significant effects.

[41]RNA (ribonucleic acid) does much of the moving and organising of DNA products in the cell; it is a very similar compound, but with a different sugar backbone (ribose rather than deoxyribose). Genes are not made of it (except in certain viruses).

Mutations can occur in any part of a gene, coding or non-coding; since DNA is so mutable, there are huge numbers of possible structures for the same gene, as well as for non-coding regions. These differences are called 'polymorphisms'. The existence of an almost astronomical number of possible polymorphisms accounts for both individual uniqueness and group similarity; and some polymorphisms are intimately involved with disease as well. Here is an example of the possible effects of a simple polymorphism that happens to affect an important gene. (This is not a mood-disorder gene, but I use it because it is about the simplest and most direct example of a complicated process.) On human chromosome 4 is a gene that codes for an as yet poorly understood protein called huntingtin. It is named after Huntington's disease (a fatal neurological affliction producing movement disorders and dementia), which results from certain alleles of this gene. Part of the gene consists of a string of glutamine (CAG) repeats. Most people have between 6 and 35 repeats, on average around 10-15. If however there are more than 39, the result is Huntington's disease. This gene is dominant, so that anyone who has one copy of a variant with too many repeats will get the disease. But the age at which the symptoms begin is governed by the number of repeats: 39 yield a 90% probability of dementia by age 75, with first appearance of symptoms about 20 years earlier; 50 will produce dementia in the late 20s.[42] If we adopt a notional minimum of six and a maximum of fifty CAG repeats, this gives us no less than 44 possible variants of this one gene with respect only to this one feature.

This illustrates in outline what one aspect of 'genetic causation' can be like; genes are highly variable, and minute changes can either predispose to or produce diseases. Everybody 'has the Huntington gene'; but only certain forms will induce the disease, and the timing of onset depends on nothing more than the number of glutamine repeats in one particular stretch.

'Genes for depression'

> A generation ago, few mental health professionals believed that inherited vulnerabilities could be central to the development of psychiatric illness. Fearing that discovery of a genetic diathesis might cast a stigma on patients and lead to therapeutic nihilism, many clinical observers found social and developmental reasons to explain the inescapable fact that mental illness runs in families.
>
> — Eliot S. Gershon, Genetics (1990)

If physical traits run in families, they are obviously 'genetic': they are encoded in the genes handed down from generation to generation. So dramatic examples like the 'Habsburg lip', haemophilia in the male descendants of Queen Victoria, Huntington's disease. Other traits run in families but seem at first sight 'environmental' rather than genetic: intelligence, musical or artistic talent, literateness, criminality, professions (lineages of bookies, doctors, musicians, gangsters ...).

In the standard lay interpretation, the shape of one's lower lip is purely a matter of genes,

---

[42] For details and references, see Ridley 1999: 55ff.

or ‘nature’; becoming a gangster in a family of gangsters is a matter of environmental influence and upbringing, or ‘nurture’. There has been a strong tendency for centuries to partition all human attributes into two classes: those due to nature and those due to nurture. This has led to one of the longest-running and most futile debates in the history of philosophy and the social sciences. It reflects a characteristic (and not in itself necessarily reprehensible) human tendency to prefer dichotomies to more complex models for just about anything. We do not, without special training, normally feel very happy with world-views that have central grey areas or ambiguities.

This tendency is supported by public lack of genetic sophistication, reinforced by sensationalist and irresponsible science journalism (‘gene for Schizophrenia/Dyslexia/ Homosexuality discovered’). Very few genes in fact are ‘for’ one particular gross character, and of course no genes are ‘for’ an illness: rather anomalies in certain genes may produce or conduce to certain illnesses. And very little of the inheritance of complex traits is strictly Mendelian anyhow, so all the complications mentioned above are likely to be involved.

Making a sharp distinction between genes and environment is also a fundamental error. Genes always operate in *some* environment or other. The genome itself is an environment: genes have to work together, often in exquisitely timed sequences. The embryo is an environment, so is the fetus; their environment is the womb; the neonate, the toddler, the older child, the adolescent, the adult are also environments; and this stack of environments is set inside another, consisting of the mother, the father, the family, the immediate conditions under which the individual is raised, the society, the culture ... Nothing is ever genes or environment, but always genes *and* environment.

The gene/environment dichotomy becomes even shakier when we observe that the same set of genes may code, under different environmental conditions, for different phenotypes—even in such an obvious matter of ‘nature’ as sex. Thus the sex of an alligator or crocodile is dependent on the position of its egg in the nest: the sex of the embryo is controlled by the temperature it matures at. A human example is the inheritance of multiple sclerosis. If one identical twin has it the other has a 30% risk of having it as well; so there must be some genetic involvement, as well as other factors, since the overall population incidence is much lower. And MS is commonest in people of northern European origin living in temperate climates. This is characteristic gene/environment interaction, with variable penetrance and expression under different environmental conditions.

Nobody is bothered by characters like eye-colour being genetically determined; but many people abhor the idea that ‘higher level’ human traits like temperament, intelligence, creativity, even propensity to mood disorder— anything ‘behavioural’ or ‘mental’ or ‘cultural’—should be. The ‘libertarian left’ (in the American sense, rather than the European sense of ‘social democrats’), in crude summary, would rather have these depend on environment alone. If they did not, so the standard argument goes, this would open the way to loss of opportunity and non-fulfilment of potential, a kind of fatalism (if it is ‘in the genes’, there is nothing to be done about it). It would encourage stereotyping, crude genetic determinism and nihilism, would erode the notions of ‘free will’ and human perfectability, and might even lead to a kind of Hitlerian

eugenics. Such hard-line 'environmentalism' is a political rather than a scientific position; but it actually makes empirical claims.[43]

Such claims are testable, because they make implicit predictions. For instance, if some trait *T* is purely environmental and not genetic, it follows that: (a) identical and fraternal twins, and non-twin siblings growing up in the same environment, should not differ significantly in incidence of *T*; and (b) twins separated at birth and raised in different families should with respect to *T* be more like their adoptive parents than their biological ones. With many features of temperament and susceptibility to mood disorder this turns out to be false. For instance:[44]

> 1. The overall population incidence ('lifetime risk') of Bipolar Disorder is 1%; that of Major Depression (recurrent or nonrecurrent) about 5%.
>
> 2. If you have a first-degree relative (parent or sibling) with Bipolar I, your lifetime risk for Bipolar Disorder (I or II) goes up to 8%, and your risk for Major Depression to 10%; that is you are 8 times more likely to suffer from Bipolar Disorder and twice as likely to suffer from Major Depression as an 'average' member of the population without this family background.
>
> 3. If you have a first-degree relative with Major Depression, your lifetime risk is 10%, again twice the population average.
>
> 4. If you have a first-degree relative with Major Depression whose first episode occurred before the age of 20, your lifetime risk is 30%, six times the population average.

These figures hold regardless of whether you share or ever have shared an environment with the relatives in question, so there must be some genetic element involved. This is even clearer when we consider the evidence from twin studies.

Twins are a rich source of information on nature vs. nurture; in particular, when pairs of twins have been separated at or near birth, and raised in different families, and good records have been kept of their fates. In many highly bureaucratised countries with state health services, like Denmark, the record keeping has been particularly good; some of our best data comes from a series of Danish twin and adoption studies (cited in Gershon 1990). These looked at

---

[43]For an eloquent, beautifully argued and entertaining destruction of the 'environmentalist' position, demonstrating how much of 'human nature' is actually hard-wired, see Pinker 2002. My own view is that after that book the nature/nurture debate is no longer necessary, but there are still a lot of people who don't feel that way.

[44] The material in this section is based largely on Gershon 1990, Souery *et al.*1997 and Barondes 1998: chapters 7-10. For a detailed introduction to the genetic epidemiology of psychiatric illness see Scham & Kendler 2008. As far as anybody knows, these distributions appear to hold for all human populations. There is nothing 'ethnic' or 'cultural' about these illnesses—though there may be in their characteristic presentations.

‘concordance’ in identical vs. fraternal twins: i.e. the statistical risk of a co-twin suffering a disorder if the other twin does. Recalling that the worldwide control population incidence is about 1% for bipolar disorder and 5% for unipolar, the following results are of interest:

1. Conflated figures (bipolar and unipolar). Identical twins: concordance 67%; fraternal twins 20%.

2. If either of a pair of identical twins is bipolar, the concordance is 79%; if one is unipolar the concordance is 54%.

3. If either of a pair of fraternal twins is bipolar, the concordance is 24%; if unipolar, 19%.

4. Breakdown of conflated bipolar concordances: If one of a pair of identical twins is BP I (the more severe type of bipolar disorder), the concordance is 80%; if one is BP II, 78%.

5. In identical twins where one is unipolar, and has had three or more episodes, the concordance is 59%; if one has had fewer than three episodes, the concordance is 33%.

The lowest concordance, in unipolar fraternal twins, is still nearly twenty times the average lifetime risk. The only reasonable explanation for this heightened risk is inheritance—but of a complex kind, involving a difficult and as yet not understood combination of polygenes, incomplete penetrance and variable expression—the most intricate possible type of inheritance.

Adoption studies and records of incidence within biological families provide striking reinforcement. The former distinguish genetic heritage from environmental influence; the latter display long-term inheritance patterns (from data summarised by Gershon):

1. In adopted children presenting with bipolar disorder, 31% of the biological parents had mood disorders, as opposed to 12% of the adoptive parents.

2. In biological families, if one parent is unipolar or bipolar, and the other normal, the overall mood disorder risk for a child is 27%.

3. In biological families, if both parents have mood disorders, and one is bipolar, the overall risk to the child is 74%.

There are a number of confounding variables here, but they do not seriously affect the general picture. One is ‘assortative mating’: the unsurprising tendency for people with mood disorders to marry other people with mood disorders. This may produce a statistical loading, since it increases the presence of whatever it is that is heritable. In the biological family data (2-3 above), there may be an environmental factor as well: if one or both parents have mood disorders, the

family environment is likely to be highly stressful, so that the children—who are at high risk of inherited vulnerability anyway—are particularly prone to having this exposed by environmental triggers. Once again, the intricate overlap of heredity and environment makes the apportionment of responsibility difficult: but the genetic element still remains central.

This suggests that at least part of what is heritable in mood disorders is not disease *per se*, but a vulnerability or predisposition to respond to stress and other environmental triggers more sensitively than others. And since the mode of inheritance seems to be so complex, the most likely explanation is that there are many genes involved, or many variants of a few genes, and perhaps no two depressives will show exactly the same pattern. There have been an enormous number of studies attempting to determine just which genes may be responsible for this vulnerability; very few have been conclusive, but one is well supported.[45]

There is a well-studied gene called 5HTT, or the Serotonin Transporter Gene (5HT is a chemical shorthand for serotonin). There is one part of this gene—the promoter region—that shows a major polymorphism. There are two basic types, called 'long' and 'short': this refers to the number of CG repeats in this region, which is tremendously variable. The presence of long or short alleles of this gene is highly predictive for a person's propensity to develop depressive illness in the presence of stress.

Since the child inherits one copy of the gene from each parent, there are three possible long/short combinations, which predict the risk for depression as follows:[46]

(i) 2 long alleles: little risk
(ii) 1 short, one long: moderate risk
(iii) 2 short: highest risk (two-thirds of these subjects show depression by their mid-20s)

Carriers of two short alleles also show considerably elevated base-line activity in the amygdala, whether depressed or not. This may be the clearest neurological marker of the effect of the two-shorts configuration: a state showing an increased likelihood of anxiety.

It is clear that this gene does not directly 'cause' depression; it codes a vulnerability to depression under stress, with clear involvement of the serotonin system and limbic structures. Possession of any particular pattern is a prediction of risk, not a condemnation to depressive illness. It is probable, judging from population distributions, twin results, etc. that there is no gene or gene combination that directly 'causes' mood disorder: rather that there are genes that make predictions of risk. But this still makes depression significantly a genetic disease.

---

[45] For a short general account of studies of particular genes and/or chromosomal regions in depression see McGuffin 2008.

[46] Caspi *et. al.* 2003. This is a 'prospective' study, i.e. it followed 1000 subjects from infancy to young adulthood. For discussion and summary http://www.psycheducation.org.mechanism/MechanismIntro.htm. As this chapter was being completed a new study appeared that argues that this effect is an artifact of the analysis, and that the long/short allele distinction does not have the claimed effect. See Risch *et al.* 2009.

Another interesting and fruitful line of enquiry has been the study of the genetics of temperament. It has been clear for a long time that basic temperamental variables like shyness vs. extraversion, anxiety vs. a sanguine temperament, novelty-seeking vs. avoidance of danger, are to a very large extent heritable.[47] Temperamental variables are certainly part of what defines vulnerability; I return to the question of temperament and personality in depressive disorder in chapter 7.

The causes of depression: summary

It is quite possible that the really ultimate causes of depression or a propensity to it lie far back in our mammalian ancestry; all mammals have similar brains, use the same neurotransmitters, and show similar responses to stress. Syndromes that look like depressive disorders, and respond to standard medications, have been observed or produced in cats, dogs, mice and nonhuman primates. And indeed it is quite easy (perhaps too easy?) to see depression as very similar to characteristic mammalian responses to aggression and loss of status and control: social withdrawal, self-concealment, 'playing possum' to avoid excess energy expenditure in situations where it is clear (or it seems clear) that nothing is to be gained by acting. Bipolar disorder is rather more complex from this point of view, though there are (admittedly speculative) arguments for the manic or hypomanic phases having a selective advantage (bipolars are often peculiarly sexy when hypomanic). But given the existence of 'mixed' manic/depressive episodes and the possibility of unipolar disorders becoming bipolar and vice versa, and uniform cross-cultural incidence of both unipolar and bipolar disorder, it is clear that something in our evolutionary endowment predisposes to these disorders, and that they are stably maintained in modern populations.[48]

So let us assume that part of our species-specific or general mammalian heritage is a certain distribution of vulnerability to mood disorder in human populations. This is a somewhat speculative notion, though I think rather well supported; but the situation and history are not really well enough understood at present to build a solid evolutionary theory. But with this possibility in mind, we can produce a reasonable summary of the causal picture as we now understand it. Drawing together the information and arguments in this and the preceding two chapters, the causal chain leading to depression has at least three major components:

(i) Panhuman programmed behaviours of one kind of another, e.g. stress-avoidance

[47] See Hamer & Copeland 1999, Kramer 1994: 184ff, Whybrow 1997: 317 for discussion and references to this work.

[48] These facts and interpretations have led to a flourishing tradition of 'Darwinian psychiatry' or evolutionary psychiatry'. For an introduction see Baron-Cohen 1997, McGuire & Troisi 1998. Schizophrenia is similarly genetically influenced and uniformly distributed, and may also be an ancient heritage: see Horrobin 2001.

mechanisms, other ancient mammalian or even general vertebrate responses and the genetic apparatus that makes them possible. We might think of these as macrohistorical or phylogenetic enabling mechanisms—the 'species substrate'.[49]

(ii) Individual genetic vulnerability. All mood disorders have a major genetic component. This is probably realised primarily in the structure of temperament, which in turn may be largely defined by innate dispositions of particular neurotransmitter and hormonal systems to respond with different degrees of intensity to environmental challenge. The important factors here would be polymorphisms involving the HPA and amine neurotransmitter systems, which together would define an individual 'reactivity style'—in practical terms a vulnerability to dysregulation under stress, or a predisposition to certain kinds of reactions or behaviour. This would be the 'individual substrate'. For instance, a highly reactive temperament, prone to 'inappropriate' response, anxiety, depression, etc. might produce a particular personal reactivity-style that then interacts with life history ('nature' vs. 'nurture'). Here we are on the border between the macrohistorical (evolutionary) and the microhistorical (individual) perspectives.

(iii) Life history. This is a loose term for all the contingencies faced by the individual, who we assume is already carrying some genetic loading for mood disorder. Everything from *in utero* factors (nutrition, oxygenation, position, birth order) through early childhood and later experience interacts with (i) and (ii). The result—in the appropriately specified individual encountering the appropriate stressors—is eventually a mood disorder. An individual life-history is made up not only of the things that happen to the individual, but his responses (e.g. hyperactivation, kindling). All things being equal, a highly reactive personality, and if this too is genetically specified, as it almost certainly is, a personality with a particular kind of temperament (e.g. hypothymic, dysthymic, cyclothymic) will be more vulnerable to certain classes of life-events than one with a different set of specifications. This inborn temperamental scaffolding fairly accurately predicts later reactivity. In combination with the self evolved by the individual in the course of a history of interactions with the world, in particular the reactions of its owner to external stressors of various kinds, and the internalisation of these stressors, it serves as a predictor of later mood disorder.

One thing we do not understand is what the predictors are of unipolar vs bipolar disorder. It would seem that the mechanisms for the depressive component are the same; but there are obviously other genetic and physiological systems involved.

[49]For some intelligent (if brief) discussion of possible evolutionary motivations for and current advantages of depression, see Keedwell 2008. Keedwell unfortunately appears to believe that most depression however is not serious, and that it should not be called a 'disease'.

This is an oversimplified account of the complex causality of depressive disorders, based on a sketch of current knowledge. As research continues we will certainly learn more; but the basic categories appear well established.