

1 The problem of fatigue

Background

Fatigue is a pervasive influence on human life, experienced by everyone on a regular basis. It may be felt as a low mood (tiredness, weariness, lethargy) or unfocused mental state (distraction, frustration, discomfort), or as an unpleasant bodily state, including headaches, tension, and vague pains in muscles and joints. It is also implicated in everyday disturbances of mood and quality of life, and, in more intense cases, can be felt as physical exhaustion, a total incapacity for any exertion, a profound lack of motivation, or depression. In terms of cognitive activities, fatigue is associated with problems of completing – or even starting – tasks, particularly where there is a requirement to sustain high levels of effort over long periods. In addition, fatigue (along with headache and colds) is among the most frequently reported health complaints in primary care clinics in Western countries, a feature of almost all illnesses, and a common after-effect of surgical intervention. Yet, fatigue remains a puzzle. How is it that we can feel tired when we do not appear to have done very much? How is it that we appear to be able to recover so quickly under some conditions, but not others? What is going on when weariness following a hard day at work can be banished by going for a run or a session at the gym? Why do some kinds of activity make us feel tired, while others, equally or even more demanding, do not? Just what is fatigue about, and how does it come to play such a significant part in mental life? Does it have an adaptive function, or is it simply an end state of the failure of the normal process of energy management? Or is it something else altogether?

Fatigue in the modern world is widely regarded as a major problem for health and wellbeing. Endemic tiredness is recognized not only in practical areas of life such as work and driving, but also in everyday experience. General practices are beset with patients reporting being ‘tired all the time’, and there is increasing clinical recognition of the related condition of chronic fatigue. Yet, despite the widespread general

interest in fatigue, and the plethora of popular books on the subject – an Amazon search on mental energy and fatigue conducted in early 2009 generated over 500 such books – the topic is poorly understood. This is true even within the scientific community, where it has been surprisingly neglected in terms of systematic scientific investigation. Following the stimulation of interest in human work and fatigue during the late nineteenth century, there was a proliferation of research on the topic, reaching its peak during the 1930s. The period between 1900 and 1940 was characterized by a wealth of detailed empirical studies, widespread theoretical interest, and a concern with application of knowledge to both educational and industrial practice. Since that time, fatigue has gradually receded from the scientific landscape, so much so that it now rarely appears even in the index of modern textbooks of cognitive psychology or even work psychology, and not at all in recent monographs and major reviews on attention and performance (Logan, 2004; Pashler, 1997; Pashler, Johnson & Ruthruff, 2001; Styles, 2006).

This trend is illustrated in Figure 1.1, which summarizes the results of a search on PsycARTICLES for journal articles that included the word ‘fatigue’ in the title, published during successive ten-year periods. The pattern is somewhat obscured by the fact that the number of academic periodicals (in all fields) has increased by a factor of several hundred during this period, from around 100 at the beginning of the twentieth century to a current total of somewhere between 20,000 and 50,000, depending on definitional criteria.¹ I could not find the relevant data for psychology and mental health, but have assumed that the growth rates are similar to those in other disciplines, as estimated by Mabe and Amin (2001), and shown here on a logarithmic scale. Figure 1.1 also indicates the very high values for fatigue publications over the two most recent decades. In fact, little of this dramatic upturn refers to the traditional laboratory research on work and fatigue that is the core of this monograph. Instead, it reflects the renewed interest over the past 20 years or so in research in two areas of practical significance. One is a concern for the effects of fatigue in the form of sleep deprivation, especially in military and transport applications. The other, a much stronger influence, is the proliferation of research on fatigue as a clinical problem within medical and neurological contexts. Such a resurgence of interest is welcome, and is likely to provide a valuable focus for fatigue research from many different directions: not only clinical, but

¹ These numbers are estimated from various sources, all based on Ulrich’s *Periodicals Central*. <http://ebookcentral.proquest.com/lib/christianalbrechts/detail.action?docID=1357321>.

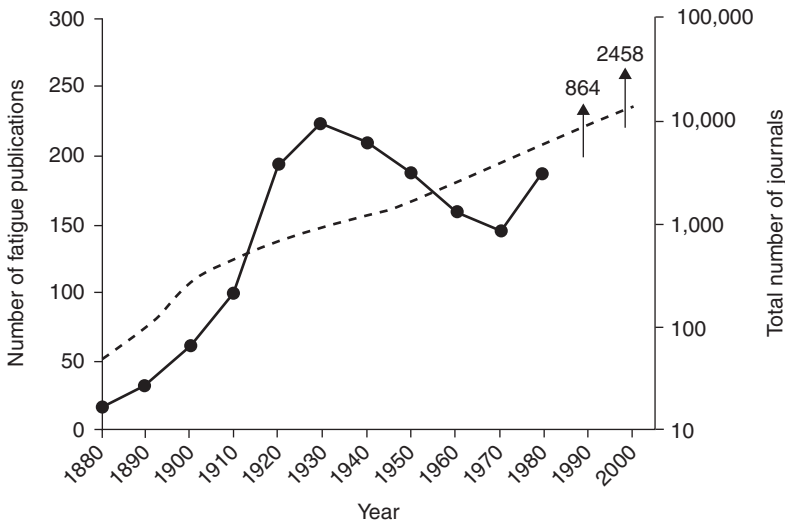


Figure 1.1 Changes in frequency of fatigue publications 1890–2000

cognitive, physical, sleep, stress and neuroscience. Even so, if an adjustment is made for publication opportunities, current publication rates of articles on fatigue are no more than around a tenth of the peak levels achieved during the 1930s and 1940s.

Of course, the tailing-off of interest in fatigue is not, in itself, noteworthy. Science is as much prone to fashion as any other aspect of human experience, and other topics take their turn in the limelight. Research on vigilance and monitoring, another topic within the attention and performance genre, emerged during the 1950s, and flourished during the second half of the twentieth century. The parent discipline of attention had an early peak, along with fatigue, in the 1920s and 1930s, before fading from view until its revival during the 1960s at the heart of the new cognitive psychology. What is strange in the case of fatigue is not that it has declined in intensity but that the problem appears to have been more or less abandoned (at least by experimental psychologists), and before it had been even partially solved. Instead, fatigue is now predominantly a topic for medical research, with a focus on its role not only in chronic fatigue but also in a wide range of other chronic illnesses. Even today, there is no widely accepted view of what fatigue is, no mature theory of its origins and function, and little understanding of the relationship between the several different types of fatigue (mental, physical, sleepiness) that are routinely confused in the scientific literature.

Rationale and focus of the book: the adaptive role of fatigue

The present monograph is the first dedicated to the scientific treatment of the topic of fatigue for more than 60 years, since Bartley and Chute's (1947) comprehensive review. Since then, publications on fatigue have been confined to edited symposia (Ackerman, 2011; Floyd & Welford, 1953), collections of chapters (DeLuca, 2005; Hancock & Desmond, 2001; Hockey, 1983; Matthews, Desmond, Hancock & Neubauer, 2012; Simonson, 1971; Simonson & Weiser, 1976), and individual chapters in textbooks or more general collections (Craig & Cooper, 1992; Hockey, 1986, 2011; Holding, 1983; Matthews, Davies, Westerman & Stammers, 2000). While many of these have usefully updated the literature and provided new ideas and insights, they have been necessarily eclectic, and have generally had little lasting impact on our understanding of what fatigue is: what causes it; what its function is (if any); what brain processes are implicated; and so on. Some of these sources also deal with the relationship between fatigue and stress, as a package of closely related issues, and it will be clear that the present book also makes no sharp distinctions between the two. Rather, they are considered to be different facets of the same adaptive process, in which the process and feeling of fatigue refer to the complex pattern of changes that follow a sustained attempt to maintain task goals under threat from environmental or task stressors. The book offers a reinterpretation of the nature of the experience of fatigue, starting with an analysis of its historical and social context. It proposes a new theory, based on the idea that such experience is a natural, adaptive feature of mental life. While fatigue is regarded as a major problem for present-day society, this was not always the case. So how did it get to be this way? And what can we do to help reduce its impact on everyday wellbeing and chronic ill-health?

The motivation for this monograph was to provide a new perspective on the nature of fatigue. At its core is the hypothesis that the experience of fatigue serves an adaptive signal function for the effective control of actions and human motivation. I argue that, rather than interfering with our ability to carry out tasks by wearing down our energy or resources, fatigue makes us aware of the opportunity costs of current activities, and of the attraction of neglected needs and alternative goals. This is a long-held perspective that has become lost from scientific view. For example, the influential Italian physiologist Angelo Mosso said of fatigue, that:

what at first sight might appear an imperfection of our body, is on the contrary one of its most marvellous perfections. The fatigue increasing more rapidly than the amount of work done saves us from the injury which lesser sensibility would involve for the organism. (Mosso, 1906, p. 156)

Such insights are also evident in the writings of Edward Thorndike, one of the most significant experimental psychologists of the early twentieth century. Thorndike (1900) rejected the idea of fatigue as a state of reduced effectiveness, likening it more to a state of mental discomfort or aversion to mental activity. He argued that:

feelings of fatigue, such as they were, were not measures of mental inability ... we can feel mentally fatigued without being so, that the feelings described above serve as a sign to us to stop working long before our actual ability to work has suffered any important decrease which an experimenting psychologist could measure and use as a warning to us. (p. 481)

A similar perspective was offered by Bartley and Chute (1947) in their comprehensive review of the problem. They interpreted the emergence of a feeling of fatigue within an individual as part of the transaction between the performer and the environment, representing a change of orientation from acceptance and engagement to one of discomfort, resistance and aversion to continuing with the present activity. The conflict between present and desired goals demands (or at least invites) a reappraisal of priorities.

While the approach and content of the present book are inevitably led by my personal research interests over the past 25 years or so, they are informed by an extensive literature. My goal in writing this book is to provide a broad context for understanding the meaning and function of fatigue, through the use of a wide range of sources: empirical and theoretical; experimental and clinical; modern and historical. I believe that a better understanding of the problem of fatigue will have benefits not only for psychological theory, but also for managing fatigue on a practical level: within work design, everyday wellbeing and mental health.

It will be clear from a quick leaf through the pages of this book that its primary focus is on the experimental psychology of *mental* or *cognitive* fatigue. Yet, the title suggests a concern with the unqualified topic of fatigue in general. This more inclusive term was chosen deliberately to reflect a desire to reconnect the various facets of the problem that have become dissociated over the past 100 years or so. These include not only fatigue from mental activity, but also issues related to sleep disturbances and physical work. My strategy, in basing the book on mental fatigue, is to emphasize the centrality of the cognitive and subjective experience of fatigue, in understanding not only mental fatigue itself

but also the impact of sleep deprivation and physical work on mental processes. This is not to say that all fatigue is mental fatigue: that there is no need to look beyond a general explanation; on the contrary, it is clear that fatigue from sleep disturbances is associated with specific needs and brain mechanisms, while physical fatigue involves muscular and metabolic demands far in excess of those met in cognitive tasks. However, I argue that, while the various forms of fatigue appear to have distinctive aetiologies, contexts and forms of expression, the development and management of mental fatigue underlies or plays a major part in all of them. A comprehensive review of the literature in these different specialist areas is not a practical goal for a monograph of this kind. Instead, I make reference to physical fatigue and sleepiness whenever it is appropriate to do so throughout, and attempt in [Chapter 7](#) to summarize the major issues relating to these alternative manifestations of the fatigue problem, and to consider how they may be integrated into a general framework.

Another deliberate focus of the book is on short-term (transient) effects of fatigue – the state experienced under conditions of acute task demands or stress, but that normally recedes when more favourable conditions prevail – rather than on enduring problems of health and wellbeing such as chronic fatigue. Inevitably, a systematic treatment of chronic fatigue is beyond the scope of this book. However, I again try to address relevant issues throughout, and, in [Chapter 8](#), review the core issues on persistent fatigue of different origins, including work and problems of chronic ill-health. I also put forward a tentative dynamic model to show how pathologies of fatigue may occur through a failure to manage the response to stress and short-term motivational conflicts.

The failure of classical fatigue theory

Despite over 100 years of research and scientific thinking about fatigue, we still have no well-developed theory of how the process works or what its function might be. A recurring theme throughout the book is that there have been two major obstacles to our understanding. The first is the irresistible tendency to think of fatigue in terms of the exhaustion of energy (feeling spent, worn out, exhausted, batteries running down). A second problem has been the near-universal tendency to consider fatigue as a negative state, and as an inevitable consequence of work. Both of these assumptions are considered in this section and found to be flawed. The alternative view, proposed in this book, argues that fatigue has an adaptive function, serving the management of motivation.

Limitations of the energy account of fatigue

The most serious failure of traditional fatigue theory has been its assumption that it is caused by a loss of energy, caused by the activity of carrying out (too much) work. This is not surprising, given the widespread influence of ideas on energy conservation and transformation on nineteenth-century thought (Rabinbach, 1990). The tendency to think about fatigue in terms of energy failure can be seen to have its origins during the Industrial Revolution, in the growing awareness of the role of energy in the work of machines, and as the limiting factor in production. Within a short time, as Anson Rabinbach (1990) shows in his analysis of the social history of fatigue, the *work*→*energy depletion* framework was being applied not only to the work of machines, but also to that of humans, and further extended to include mental activity as well as the more directly comparable physical work. The idea had an immediate and major impact on both the scientific literature and everyday language, which began to use the terms *fatigue* and *energy* together in relation to human activity from around 1870. It was a short step from this to ‘explain’ the problem of fatigue – the failure to continue to produce output – as the result of a loss of energy. The feelings of weariness during and after periods of demanding work appeared to make sense within this new way of thinking about the body. I shall discuss the derivation of these ideas in Chapter 2, and argue that the roots of the link between work, energy and fatigue are to be found in the dramatic changes in working life during the second half of the nineteenth century.

Energy-based explanations are ubiquitous within psychological theorizing on motivation, personality and cognitive psychology, often taking the form of drives or resources, though they probably had their greatest impact through Freud’s psychodynamic model, which posited ‘psychic energy’ as the basic driving force of mental life (Strachey, 1953). Reinforced by followers such as Jung, Adler and Klein, such views had a major influence on cultural thought during the first half of the twentieth century, and led to the widespread acceptance of mental energy as the prime mover of motivation and action, and lack of energy as the reason why things were not (or could not be) done. The late nineteenth-century energy view of fatigue has been highly resistant to change, even in the face of the growth of psychology as a science and the emergence of new understanding of brain and body mechanisms. While (up to a point) an energy explanation can be made to work for physical exercise, it has not stood the test of time as a way of

understanding patterns of decrement in mental work. Of course, at the most basic level, energy transformations are the basis of all neural events and all behaviour. But this does not mean that it is necessarily appropriate for high level explanations of behavioural change. It is possible (though still not confirmed) that the brain uses up more glucose and oxygen to carry out more demanding tasks than simpler ones, but the difference appears to be very small, and there are never signs of anything approaching a state of a *depletion* of energy reserves (Gibson, 2007; Raichle & Mintun, 2006).

Despite the lack of any convincing role for energy variations in behaviour, we talk freely about psychological energy (or vigour, or vitality) as a primary agent of action in many areas of mental life: for example, motivation (Ryan & Deci, 2008), work engagement (Shirom, 2003), self-control (Baumeister, Vohs & Tice, 2007), or mood (Matthews, 2011; Thayer, 1989). What we refer to as ‘mental energy’ seems to be a convincing characteristic of people, readily understood as a feeling of liveliness and active enthusiasm for the pursuit of goals. Yet, there is no evidence that healthy, highly energetic individuals differ from others in terms of actual brain energy (Lieberman, 2006). If mental energy is not a function of differences in physical energy, what is it? The continued use of the energy metaphor makes it difficult to consider other explanations for what lies behind such a state.

Fatigue is not (just) a negative state

A second, related, reason for the failure of the fatigue construct to attain scientific maturity is its narrow interpretation as an *unwanted* by-product of (physical and mental) work. This is a natural consequence of identifying human transactions with the environment with the work done by industrial machines, and the exhaustion of energy in the execution of that work. However, the ensuing emphasis on fatigue as a negative state associated with the depletion of energy has not been as useful an idea as was initially assumed. I would go further and argue that such views have impeded progress towards a genuine theory of fatigue. Examination of archival material in Chapter 2 suggests that fatigue may have been experienced quite differently in the pre-modern era (before the full impact of industrialization). I shall argue that the widespread use of energy and fatigue terms in everyday language dates from the increasing shift to factory-based work over the second half of the nineteenth century; before that time, as Rabinbach (1990) argues, there is little to indicate that fatigue was experienced as a generally negative state. I will argue that the same principle applies today, even

though the environmental and cultural context of working life has changed considerably; work is not always experienced as tiring.

As others have commented (notably, Bartley & Chute, 1947), the energy depletion perspective has almost certainly been a source of distraction in the search for an understanding of fatigue. Fatigue has always been associated with work, specifically with the depletion of energy through excessive work (what I refer to as the work–fatigue hypothesis). Such a view has long been considered an inadequate account of the varied phenomena of fatigue. Bartley and Chute (1947) concluded categorically that, even for physical work, unless extreme, fatigue represented not an inability to do work, but rather a *lack of desire*: an ‘attempt to retreat or escape from a situation’ (p. 53). I examine the relationship between work and fatigue in Chapters 3, 4 and 5, in relation to stress and coping strategies, and try to determine the boundary conditions for the assumed causal nature of the work–fatigue relationship. At one level, work may represent simply the physical and mental load on the body and brain defined by the energy requirement of the activities carried out. It becomes clear that this is not a good predictor of fatigue and its consequences, except in the case of extreme physical work. Instead, the evidence shows that it matters how activities are understood by the performer: whether they are externally driven (tasks) or self-selected (leisure); or, if they are tasks, whether they allow a high level of control (discretion in how and when they are carried out). As has long been known (Thorndike, 1900), when activities are self-initiated (and sustained by personal goals) mental work is not generally perceived as tiring, and may even have the opposite effect: that of energizing the performer, as in the experience of ‘flow’ described by Csikszentmihalyi (1977, 1990; Demerouti, 2006).

Holding (1983) identified the core effect of being fatigued as an aversion to further activities that demanded high levels of effort. It is important to acknowledge at the outset the explicit link between fatigue and effort. Effort is assumed to act as a precursor to fatigue, which then triggers control activity that serves to reduce effort. In many contexts, where the time course of feelings is not well defined, the two may be almost interchangeable; phenomenologically, needing to make more and more effort is functionally equivalent to becoming more and more tired. Bartley and Chute (1947) argued that fatigue (or effort?) may be considered both a warning of the need to escape and a marker of the distress felt when this action is not carried out. They emphasized the importance of the ‘stance’ or ‘attitude’ an individual takes to the task: whether it is desired by the performer, or not; whether it leads to

engagement or resistance. This reinterpretation of the nature of the experience of fatigue has been largely forgotten, or at least overlooked in modern treatments of the problem, though it is rekindled by recent work that shows the value of considering effort in terms of its cost for behavioural direction and decision-making (Boksem & Tops, 2008; Kool, McGuire, Rosen & Botvinick, 2010; Kurzban, Duckworth, Kable & Myers, *in press*).

A motivational perspective

In contrast to the prevailing view of fatigue as a failure of energy, the approach taken in this volume is to regard fatigue as primarily affecting the selection and control of goals. As with all organisms, humans are in a state of constant dilemma between the choice of maintaining current goals and behavioural directions and switching to new ones whenever they offer greater potential benefits. A more general motivational context is the conflict between the need to exploit established sources of reward and explore the environment for new opportunities. This is a well-established principle in evolutionary biology (e.g., Tooby & Cosmides, 2005), where, for example, foraging behaviour is shown to accurately reflect changing utilities of available food sources; animals decide almost optimally whether to stay or to shift. In human behaviour such ideas are less well developed, though they are implicitly understood in theories of motivation and cognitive control (Dreisbach, 2006). In Chapter 4, I outline the case for considering fatigue as an emotion, with the adaptive function of maintaining this motivational balance. Interruptions of current behaviour allow alternative options for the control of behaviour to be entertained. By interrupting ongoing activity, fatigue provokes a reappraisal of the benefits and costs of current goals, and allows alternatives to compete for access to motivational control. As I shall discuss in Chapter 5, goals need to be protected from such intrusion only when they rely heavily on top-down executive control. Specifically in the context of work and fatigue, the act of carrying out work is assumed to be fatiguing only when it takes the form of a *task*, a goal that is driven by external or internal targets, whether for someone else or for oneself. Fatigue develops if the performer is motivated to maintain the task goal in the face of a desire to stop or change to something else, and needs to employ a high level of effort to do so.

Over a century ago Thorndike (1900) interpreted the development of fatigue as the problem of doing the right thing, rather than doing too much. The same point was being made by Cattell (1941), who argued for a strong guiding role of purpose (or goal) in preventing fatigue-related

decrement: 'There is no fatigue as long as a purpose itself is not fatigued' (p. 624). The broader view of Bartley (1943) was that: 'The basis of fatigue is conflict and frustration. One of the first significant outcomes of conflict is a sense of *discomfort, danger or failure*' (p. 161). And the idea of fatigue as a result of conflict of goals is reiterated in Bartley and Chute's (1947) volume; they concluded that fatigue was a result of conflict between competing behavioural tendencies: between doing and not doing; between doing one thing and doing another; between doing one thing for a long time or switching to something else.

The idea that the resolution of conflict is an effective basis for the control of action is a familiar one (Berlyne, 1960; Botvinick, Braver, Barch, Carter & Cohen, 2001; Norman & Shallice, 1986), with cognitive control acting to select and promote goals, maintain selected goals, and prevent disruption by competing goals. Adaptive goal-directed behaviour requires the monitoring of current actions and outcomes, and making appropriate modifications in relation to costs and benefits of different activities. As I have already stated, fatigue is interpreted here as an adaptive state, supporting the effective overall (system-wide) management of goals. In this conceptualization, the subjective experience of fatigue arises through conflict between current and competing goals or action tendencies. In effect, the feeling of fatigue is assumed to have a signal value for motivational control, providing a mechanism for resolving conflicts between current goals and other possible or desired course of actions. This approach is developed within the rest of the book by considering the boundary conditions for the experience and impact of fatigue, especially in relation to work and the performance of tasks. However, the focus is necessarily broader than fatigue itself, since fatigue is but one aspect of the complex set of control systems that manage goal activity in the service of motivational requirements. Thus, the book is also about effort, stress, coping and motivation.

Defining the field: what is fatigue?

How can we define fatigue? What are the psychological criteria for it? What does it do? These questions have recurred throughout the history of the problem, and we are still unable to provide definitive answers. There are three persisting problems in addressing the question of definition. One is the widespread assumption that it is primarily about decrement; for example, Simonson (1971) identifies fatigue with 'all processes resulting in a decrement of capacities'. The second is that, while fatigue may be recognized as being about not only decrement, but also subjective experience or even physiological changes,

different researchers have routinely given different weight to these various aspects, emphasizing one to the near exclusion of others. The third is that the term appears to have quite different institutionalized meanings across different scientific groups: the medical profession thinks of fatigue as a debilitating consequence of a wide range of diseases and systemic conditions; neurophysiologists focus on the loss of motor control leading to reduced muscular strength; exercise physiologists talk of fatigue as a limiting factor in physical performance and endurance; and sleep researchers regard it as a condition brought on by impaired sleep quality or duration. And how do psychologists think of fatigue? Throughout its history, there has been a tendency towards overgeneralization; in the words of one prominent early researcher:

Fatigue is a comprehensive term which in its widest application embraces all those immediate and temporary changes, whether of a functional or organic character, which take place within an organism or any of its constituent parts as a direct result of its own exertions, and which tend to interfere with or inhibit the organism's further activities. (Ash, 1914, p. 1)

On such a basis, it would appear difficult *not* to study fatigue, whatever one's research focus, as long as the organism does something! The problem has also been one of circularity; fatigue was often assumed simply because of the nature of the testing conditions, or the ways in which behaviour was assessed. In a frequently cited report to the Industrial Fatigue Board, the Australian industrial psychologist Bernard Muscio (1921) concluded that it was not possible to define what fatigue was *independently of the tests used to measure it*, and recommended that the term be avoided altogether. He was certainly correct in criticizing the circularity of reasoning that pervaded psychological research at that time – and not only in work on fatigue – though his conclusions are unduly pessimistic.

Ideally, what is needed is to show a correspondence between an acknowledged state of fatigue and various consequences of that state. In early research effects were typically expected as a reduction of scores on specific tests, but these tests were also often used to define what was meant by the state itself (or the characteristic expression of it): 'a condition caused by activity in which the capacity for repeating the activity that caused it is diminished' (Muscio, 1921, p. 35). Muscio argued that there needed to be an independent measurement of the fatigue state, such as a set of physiological changes. This is what the early fatigue researchers had anticipated they would find in simple muscular or metabolic measures of energy loss, though they were uniformly unsuccessful. In fact, circularity of this kind is still frequently found

in journal articles, again concerning not only fatigue, but also stress, arousal, effort, and many other motivational constructs. These all have in common the difficulty of finding any unequivocal bodily indicators of what are broad, complex states. I would argue that, while Muscio's analysis reflects a realistic appraisal of the state of the art at the time, his inference is flawed. How is fatigue different in this respect from other psychological states that rely on introspective reports as the primary evidence for their occurrence? What about effort? Anxiety? Pain? In some cases, there may be the possibility of measuring concomitant physiological changes, but these are not completely reliable, nor uniquely attributable to changes in the relevant state. The essence of fatigue (as of effort, anxiety and pain) is not its physiology or its effect on performance, but its undeniable subjective quality; the feeling of mental tiredness is one that is universally recognized and understood.

A working definition

It is, however, necessary to be clear about what we do mean by fatigue: what it is and what it is not. For this purpose, a good starting point is provided by the criteria suggested by Bartlett, in the Floyd and Welford (1953) symposium on fatigue:

Fatigue is a term used to cover all those determinable changes in the expression of an activity that can be traced to the continuing exercise of that activity under its normal operational conditions, and that can shown to lead, either immediately or after delay, to deterioration in the expression of that activity, or, more simply, to results within that activity that are not wanted. (Bartlett, 1953, p. 1)

Bartlett's definition has a number of features that are worth emphasizing. First, fatigue is identified as a *process* – a growing problem associated with *continued* activity of a task. This remains the core definition of fatigue effects in task performance, especially when a person has been carrying out a highly demanding task or dealing with stressful events. Second, like Bartley and Chute (1947), Bartlett is careful not to identify effects of fatigue with decrement per se. He suggested that 'feeling tired' (with its associated signs of physical discomfort) may, in fact, represent a somewhat late stage in the fatigue process, 'when a good many unwanted effects have already invaded performance', and identified three phases in the development of fatigue with repeated work, before any discernible overall reduction in work output or speed occurred: (1) loss of timing and control of successive task elements; (2) loss of organization, or adjustments in the way the task is being managed; and,

finally, (3) feelings of fatigue and physical discomfort. Although performance decrements are generally regarded as the gold standard of fatigue research, as I show in [Chapter 3](#) they are by no means routinely observed. In [Chapters 4](#) and [5](#) I argue that this is because performance may be protected by the use of effortful strategies that help maintain task goals, especially when individuals are highly motivated (Hockey, 1997). However, the act of performance protection under increasing effort is expected to result in the emergence of the end state identified by Bartlett's definition, a delayed effect of fatigue on performance. Although he assumes a gradual breakdown of performance, the development of fatigue may also have an impact on post-task activities, what I refer to as the fatigue after-effect. In addition to a gradual reduction in commitment to the continuing task, fatigue inculcates a preference for low effort strategies in the period following a demanding or stressful work session. Following its clarification by Broadbent (1979), Cohen (1980) and Holding (1983), this type of effect has become increasingly studied and identified as a more typical consequence of fatigue, particularly in realistic work situations, or when participants are highly motivated to perform well on primary tasks (Hockey & Earle, 2006; van der Linden, 2011).

I shall retain the essence of Bartlett's criteria to define the range of phenomena to be explained, with the addition of an index of the costs of effort. The use of compensatory effort to account for the discrepancy between fatigue and impairment was understood by many early researchers (e.g., Bitterman, 1944; Freeman, 1931; Ryan, 1944; Thorndike, 1900) and provides a valuable reference point for interpreting fatigue states; an absence of impairment, when it is accompanied by higher levels of effort and tiredness, is an indicator of the increased difficulty of maintaining task goals, a defining feature of what is meant by fatigue. The feeling of fatigue may be regarded as a direct result of the use of increased effort to maintain task goals and protect performance during periods of demanding work. In some conditions effort has to be increased much more to prevent major impairment in the task (what I refer to as the strain state). Putting this all together ([Table 1.1](#)) it is clear that the patterning of fatigue may take three general forms: acceptance of interruption and its consequences; normal (manageable) resistance to interruption; and a state of strain, where extended resistance (and sustained effort) develops into an aversive state. Note that performance decrement and fatigue feelings are not considered to co-occur in the 'pure' states, though most responses to tasks involve a mixture of the two. Under strain, high effort may not be enough to sustain task fidelity and both effects may be observed.

Table 1.1 *Possible patterns of the fatigue response to work tasks*

<i>Fatigue mode</i>	Subjective state	Performance decrement	After-effects
<i>Acceptance</i>	minor (transient) fatigue	yes	no
<i>Resistance</i>	increasing effort and fatigue	no	yes
<i>Strain</i>	sustained high effort and fatigue	slight	yes

Fatigue and related feelings

Fatigue appears to be a rather distinctive state, and, in later chapters, I develop the idea of fatigue having the core properties of a basic emotion, with attendant implications for goal management and behavioural expression. However, it appears to act as a highly generalized background emotion, with figural detail being provided by other more specific states, such as the anger and distress of young children (Tomkins, 1963). Because of this it has conceptual and phenomenological overlaps with many other motivational and emotional experiences: notably boredom, depression and anxiety. Fatigue is also functionally related to the experience of effort, and is a natural outcome of the stress response, as well as being a commonly reported symptom of many illness states. These associations with other feelings sometimes make it difficult to identify fatigue as the key factor in particular situations, and may influence the kind of explanation of behaviour that seems appropriate. In other cases it serves to remind us that states such as fatigue are nearly always more complex than typically assumed. Nevertheless, there is a need for clarity about the relation of fatigue to these other feelings.

Boredom. The link with boredom is particularly strong, and has been a source of ambiguity since the 1930s. Myers (1937) differentiated between mental fatigue and boredom, though recognized that they may have similar effects on performance. Specifically, he defined boredom as ‘the outcome of a failure to find interests which can maintain spontaneous or voluntary attention’ (p. 298), while fatigue was ‘a general impotence to concentrate attention and to act purposefully, intelligently and creatively’ (p. 299). Myers’ main distinction, then, is that boredom is a failure to attend to a specific source of information, while fatigue is a more general failure of concentration. Based on an extensive programme of work, Barmack (1937) argued that boredom

Copyright © 2013. Cambridge University Press. All rights reserved.

developed in task situations under low intrinsic motivation, and was a state of conflict between remaining in the situation and wanting to get away from it. This conflict was partially resolved by the onset of a state of detachment or sleep-like withdrawal, associated with reduced sympathetic activity. Stated in this way, boredom looks remarkably like fatigue, as I have defined it here. Berman (1939) preferred to think of boredom as satiation, a general response to continued stimulation of all kinds, but without the need to invoke a mental state. Both boredom and satiation are clearly related to fatigue, though without the sense of task striving; the most basic of the three is satiation, since it can occur in the absence of a task, while boredom may be thought of as fatigue without the effort to remain actively engaged with task goals.

Welford (1968) considered the relation between fatigue and boredom in terms of the arousal-information processing theories of the day. The construct of general arousal became very popular in applied experimental psychology during the 1960s as a framework for accounting for effects of stressors on performance, through its supposed 'inverted U' relationship with performance (e.g., Broadbent, 1963). In relation to human performance, such views assumed that general arousal was dependent on the level of information in task events, and that there was an optimal level of arousal for effective performance. Welford suggested that fatigue occurred as a consequence of information overload (too high a level of arousal), and considered decrements under fatigue to be caused by the consequent increase in 'neural noise' under the sustained stimulation of demanding tasks. This seems counter-intuitive, since a state of 'over-arousal' is usually taken to mean one of stress or panic, rather than the more familiar understanding of fatigue as an inhibited, withdrawn or inactive condition. However, he did not consider the possibility put forward later in this book of the one state leading to the other, through active inhibition. By contrast, boredom was assumed to be caused by underload, too low a level of arousal, brought about by inadequate environmental stimulation; although task events attract attention they do not contain enough information to maintain active engagement.

Despite this distinction, Welford nevertheless recognized that fatigue and boredom were not always easy to separate empirically, and that both may occur in a prolonged task situation (where, for example, parts of the task are highly demanding while overall throughput of information is low). There are some contextual differences between fatigue and boredom; O'Hanlon (1981) has suggested that boredom is a response specific to situations that are highly repetitive or monotonous, not

standard features of the typical fatigue-inducing task. One possibility is that fatigue and boredom represent subjective responses to different features of the task situation. In their analysis of the characteristics of the fatigue process, Bartley and Chute (1947) argued that the aversive feelings experienced under boredom were a response to the monotony of the environment, whereas fatigue had its origins in the individual's perceived inadequacy to manage the task. Overall, however, these two feeling states are very close together, and I would argue that boredom may be a component of the broader fatigue process. In this case, the monotony of the task environment may be considered a special (or limiting) case of the demanding work situation associated with fatigue. In both paradigms, individuals are required to remain in the task situation imposed from outside, and, in both cases, they are motivated to change the environment or task, or to break out of the situation and do something else.

Anxiety and depression. A second potential source of confusion is the frequent association observed between fatigue and other negative mood states, such as anxiety and depression. Such feelings have long been recognized as occurring together in what were categorized as 'nervous disorders' in eighteenth-century England (Porter, 2001; Sicherman, 1977; Wessely, 1991), and are regularly found as a cluster of symptoms in mood disturbances. In modern analyses of mood, the relationship between them can be interpreted in a number of ways. Standard circumplex models (e.g., Russell, 1980) typically illustrate a two-dimensional solution for the underlying structure of feelings, based on factor analyses of the correlations between mood reports. This identifies mood terms varying in terms of *valence* or hedonic tone (negative to positive) and *intensity* or arousal (low to high). In this kind of analysis, feelings such as fatigued, anxious and depressed (the 'negative mood' cluster) are grouped together as negative valence, in contrast to positive mood terms such as energetic, contented and elated. The intensity dimension serves to distinguish active or aroused states (energetic, anxious, elated) from less active ones (fatigued, depressed, contented).

An alternative methodology involves rotating the axes of the factor analytic solution to give dimensions based on the diagonals of the standard solution. The best-known of these models (Watson & Tellegen, 1985; Watson, Wiese, Vaidya & Tellegen, 1999) refers to these new dimensions by the somewhat confusing labels of *positive affect* (PA) and *negative affect* (NA), both ranging from low to high. High PA includes feelings such as enthusiastic, elated and energetic, and low PA feelings of depressed state and fatigue; high NA refers to

a state of active distress, including feeling anxious and angry; and low NA to feelings such as calm and contented. Finally, a modification of the rotated model (Thayer, 1989) expresses moods in terms of two patterns of arousal: *energetic arousal* and *tense arousal*. These can be seen to be broadly equivalent to PA and NA, but with less emotional content. While the two kinds of solution are statistically equivalent, the PA/NA structure of affect has been of greater value for measuring response to stress or task demands (Hockey, Maule, Clough & Bdzola, 2000; Warr, 1990). Watson et al. (1999) have shown that they are also highly correlated with the ‘big two’ personality dimensions, extraversion and neuroticism, and correspond in affective terms to the two fundamental motivational systems associated with goal pursuit/approach and withdrawal/avoidance behaviour. I will say more about this in Chapter 4 in relation to the metacognitive role of feelings in goal management.

It makes sense on other grounds to distinguish anxiety and fatigue, as they clearly represent different mechanisms and subjective states. However, they are bound together as components of the response to stress and high demands. In a number of studies, while we have found it useful to distinguish fatigue (low energetic arousal or low PA) from anxiety (high tense arousal or high NA) as separate indicators of strain, they often occur in combination as the classic strain pattern associated with an effortful response to demanding work (Hockey, Payne & Rick, 1996). In the compensatory control model (Hockey, 1997) anxiety is identified with the response to a perceived threat from the environment. It will always occur in task situations when demands are unexpectedly hard to manage. Anxiety is also the typical precursor of fatigue in such situations, but only when sustained effort is recruited to meet these demands.

The conceptual status of fatigue

Following the early enthusiasm for research on fatigue, a general scepticism developed, expressing concern that fatigue was an unnecessary concept. This was articulated most forcibly by Muscio (1921), who argued that there was no way of defining fatigue independently of the measurements made to identify it. These problems have not been resolved and, even today, considerable confusion remains about its conceptual status. Does fatigue have a real explanatory value? Does it help us to understand the different patterns of behaviour observed under various work conditions? Does it relate to some real physical entity such

as a brain process or bodily condition? Or does it simply add an unnecessary layer of description to the empirical relationships that we observe between independent and dependent variables? Of course, these are questions that can be (and have been) asked of almost all explanatory concepts in psychology. Part of the problem is that researchers often use conceptual terms loosely, or may be unclear about how they mean them to be understood.

Such issues were treated formally in an influential paper by MacCorquodale and Meehl (1948), mainly addressed to the confusion over the meaning of explanatory terms between the competing dominant behaviourist theories of the time (those of Hull, Tolman and Skinner). MacCorquodale and Meehl distinguished between two kinds of theoretical uses of unobserved variables that mediate between or link stimuli/inputs and responses/outputs: hypothetical constructs and intervening variables. Intervening variables are abstract concepts employed to summarize the observed relationships in a way that aids meaning (for example, through their association with everyday language), though they are not presumed to have any physical reality. By contrast, hypothetical constructs are assumed to have an underlying basis in physical processes. They have properties and implications that have not necessarily been directly observed, and therefore have a stronger role in theory development, helping to generate hypotheses about hitherto untested relationships, and improving its predictive power.

One problem is that, depending on the context, the same term may be used either as an intervening variable or as a hypothetical construct. Consider *learning* as an example. It cannot be observed directly, though we believe that it corresponds to some physical activity (changes in neuronal structures) that underlies observed increments in knowledge and skill. Used in this way learning has the status of a hypothetical construct. However, in some uses, it may have the status only of an intervening variable. Suppose we observed an unexpected improvement in the performance on a well-practised motor skill between two task sessions, and ‘explained’ this as being due to learning having taken place. Here, learning is being used as an intervening variable; it adds nothing to the understanding of the observed relationship. It is also misleading, since we have not ruled out other potential causes of skill improvement in the second session, such as changes in motivational factors, effort, or demand characteristics of the task. This may seem a trivial example, since we can recognize the attribution of learning in the latter case as one of slipshod methodology; we need to control for

these other competing hypotheses before we can support learning as the 'real' cause. However, the difference is not always clear-cut and may be a matter of degree. In the case of learning we may be fairly sure of our physical ground, but what of fatigue, or of goal, energy and effort? Soames-Job and Dalziel (2001) were unequivocal in their definition of fatigue as a hypothetical construct, based on an underlying energy limitation of some kind:

Fatigue refers to the state of an organism's muscles or viscera, or central nervous system, in which prior physical activity and/or mental processing, in the absence of sufficient rest, results in insufficient cellular capacity or system-wide energy to maintain the original level of activity and/or processing by using normal resources. (p. 469)

This is the traditional view of fatigue, and one that I will argue against in later chapters, though it clearly imbues the idea of fatigue with some substance. More importantly, at this stage, is the question of whether a case may be made for considering fatigue as an intervening variable only. After all, it is most typically used as a ready explanation for performance on a task falling off over time: a convenient label for the observed relationship between a set of conditions and an outcome. It is also sometimes used to refer to the experimental condition of an independent variable (the 'fatigue' condition), pre-empting its role as an explanatory variable in the understanding of the observed empirical relationship. Because of these tendencies, Muscio (1921) cannot be strongly criticized for dismissing fatigue – in many cases, the term appeared to do no more than this – even though his argument could have been applied equally to many psychological concepts at that time. In other situations, however, and with the growth of empirical evidence, it is clear that fatigue implies a more elaborate central process. In this form, time on task is but one of the ways of inducing fatigue, and an observed decrement but one of its possible manifestations. Fatigue may also be induced by stressors and high effort, and reduced by control opportunities. It may be measured not only as subjective tiredness, but also as interruptions of response timing or after-effects of resistance to further effort. Here, fatigue is the conceptual anchor that is needed to draw together the various separate effects. Rather than have independent explanations for each we may propose a general theory that goes beyond the observable data of specific manipulations; it can account for apparently incompatible findings, and simplifies the explanatory framework. As with more widely accepted constructs, such as memory, metabolism or multiple sclerosis, or emotions such as anxiety and anger, a theory that makes use of a hypothetical construct

has to show that it can go beyond the empirical relationships observed in experiments, and carries with it a duty to predict new cause–effect patterns. On the view proposed here, I would argue that the manifestations of fatigue are (at least in principle) observable as components of a brain state. In the second half of the book, I will provide evidence for such a claim.

The key point here is to recognize that the subjective feeling of fatigue is the primary marker of the state (as is true also for anxiety, pain, hunger and other states), rather than performance decrement. The failure to find the holy grail of an objective test of fatigue is irrelevant, since it has been driven by the misguided search for evidence of a draining of central energy, leading inexorably to a waning of output or behavioural intensity. As Bartley (1943) points out:

(the) subject's report of fatigue is its *sine qua non*. One is not tired till he knows it. Prior to this one is only impaired, the realization of which becomes a component of the resulting total fatigue-syndrome. Fatigue is the desire to quit a given activity and turn to something else, arising out of the discomfort or impairment involved in pursuing the task, or in the relative failure in doing so for any reason; or it is the feeling of aversion toward instituting a given activity owing to the anticipated discomfort or relative inability in performing it. (p. 161)

However, I believe that fatigue is a more complex process than both impairment and the emergent feeling. Rather, it has subtle effects on performance even at the earliest stage of task engagement, before both of these are evident. I shall develop this idea throughout the book. In Chapter 4 I put forward an argument for treating fatigue as having a basis in emotion, like anxiety and depression. As with both of these, fatigue is both a process and a state. As with all emotions (Izard, 2009; Nesse, 1990) the process (fatigue, anxiety, depression) is triggered by an automatic response to significant environmental events, leading to a set of changes designed to resolve the problem. However, an extended or unresolved process may lead to the distinctive end state experienced as fatigued, anxious or depressed.

Some questions for a scientific theory of fatigue

In deciding upon the approach and contents of this book, I have had to consider what issues a modern treatment of fatigue should address. To a certain extent this will, inevitably, reflect my own interests and biases. However, I have tried to be as broad and impartial as possible in my choice of topics. The following is a set of issues that I believe are central

to the process of developing a scientific theory of fatigue. I have not tried to justify them here; rather, they may be seen as general postulates for the theory I develop later in the book.

- (1) *Function of fatigue.* Does fatigue have a function or purpose, for example in terms of regulating bodily or mental events? What is it for? What does it do? Early researchers considered fatigue not as an end point of an energy transformation process, but as a warning or indicator of the need to stop now, before irreparable damage occurs. This makes sense within an evolutionary framework. I explore the view that fatigue may be considered, like anxiety, to have an adaptive function, serving to protect the organism from over-commitment to specific goals, in the service of a balanced motivational strategy. By acting as a signal for rest or change it allows a reappraisal of competing needs and their values.
- (2) *Task performance.* The core empirical work on fatigue has been concerned with its role in task performance; with its assumed relationship with work and sustained attention and high workload. Such effects are known to be moderated by work conditions such as control and effort, though these have not been included in most theories of performance decrement. In general, a theory of fatigue will need to be able to account for the variability of performance under a wide range of tasks and conditions.
- (3) *Stress and effort.* The effects of fatigue are intimately bound up with the response to stress, and to the use of sustained effort as a control strategy. These effects may explain the after-effects of fatigue and be the origin of longer-lasting problems of fatigue. A theory of fatigue will need to be consistent with what is known about the stress response.
- (4) *Task goals, personal goals and control.* One of the building blocks of my approach is that fatigue is a way of preventing rigidity of goal maintenance. This is assumed to be a problem only for work that is imposed or understood as a task. For personal goals or interests (with high control or intrinsic motivation) fatigue does not appear to be a problem. The role of control is therefore central to the dynamics of the fatigue process.
- (5) *Brain mechanisms of fatigue and effort.* Are there identifiable brain processes that relate to the phenomena of fatigue and to the effort-fatigue relationship? I make the case that a strain on executive functioning is the core problem; are there mechanisms that reflect this? And can we say anything about possible neurochemical factors that may support these?

- (6) *Energy*. The traditional link between work, energy and fatigue is one of the starting points for the book. While it now seems unlikely that fatigue is caused by marked changes in glucose metabolism in either brain or body, the possibility of more subtle influences needs to be considered. Energy transactions are also implicated in the related problems of sleep and physical work, raising further questions of how feelings of fatigue and effort are related to changes in energy.
- (7) *Varieties of fatigue*. A long-running question is whether fatigue is one thing or many; whether mental fatigue, physical fatigue and sleepiness are the same thing, or whether they represent quite different processes. At this stage the evidence is not sufficient to draw unequivocal conclusions, though some commonality is likely. A theory will need to be clear about where it draws its boundaries.
- (8) *Malfunctions of fatigue*. While the normal mechanisms of fatigue impose brief disruptions on behaviour and early recovery, fatigue may sometimes persist for an evening or a few days; sometimes for weeks or months. A major question is whether this persistent or chronic fatigue is the same process without its recovery phase, or whether it is something different, with some of the same feelings. While this is not an issue that a theory of normal fatigue needs to be overly concerned about, an all-embracing perspective should be able to at least consider such issues.
- (9) *Centrality of mental fatigue*. Whatever the outcome of arguments about the plurality of fatigue states, mental fatigue is hypothesized as being, at least, a final common path. It reflects a subjective assessment of whether some activity needs to be stopped (or changed), and may also have a primary role in their management: when sleep must be resisted or yielded to; when physical endurance can no longer be tolerated.

Plan of the book

Following this introduction, [Chapter 2](#) adopts a historical perspective to consider changes in the meaning and impact of fatigue, both as an everyday term and in its scientific usage. In [Chapters 3, 4 and 5](#), I examine the core human performance issues of the work–fatigue relationship in the context of sustained work, workload, vigilance, stress and coping. This leads in [Chapter 6](#) to development of the motivational control theory of fatigue, which combines approaches from control theory with emerging knowledge about brain mechanisms of executive

function and effort. In [Chapters 7 and 8](#), I consider constraints placed on this approach by evidence from related fields, concerned with brain energy, physical work and sleep ([Chapter 7](#)) and pathological states such as chronic fatigue ([Chapter 8](#)). Finally, in [Chapter 9](#), I suggest a broad agenda for research on fatigue and its application to real-world problems.