

Book Reviews

Purple haze: the puzzle of consciousness

JOSEPH LEVINE

New York: Oxford University Press, 2001

Joseph Levine is deeply puzzled by phenomenal consciousness, and in this book he explains why. On the one hand, he is convinced that phenomenal consciousness must be a purely physical phenomenon, since only if it is so regarded can it plausibly be credited with the causal role in the production of physical effects that we ordinarily take it to possess (Chapter 1). And he rejects even the most sophisticated of recent conceivability arguments intended to show that phenomenal consciousness is not physical (Chapter 2). On the other hand, he finds it entirely baffling how phenomenal consciousness could be a purely physical phenomenon (Chapters 3, 4, 5, and 6). But he is not one of those philosophers who officially find phenomenal consciousness baffling, but who unofficially find it non-physical. For at the end of the book he claims that “the problem goes even deeper than materialism” (p. 173), and that “it isn’t really physicality that presents the problem” (p. 177).

So what does? Despite the exemplary candor and straightforwardness of Levine’s writing, the answer to this question never became perfectly clear to me. But the reason for this lies in the subject matter: Levine strives to articulate intuitive disquiets that, for all their obscurity, are very influential psychologically in discouraging philosophers from physicalism about phenomenal consciousness. He is right, however, to make the attempt, and we should be grateful to him for it, even if we go on to carp at its possible deficiencies. Levine succeeds wonderfully in bringing out just how much theorizing remains to be done by physicalists in regard to our introspective thoughts about phenomenal consciousness. Physicalists have typically offered just enough of an account of

introspective thoughts about phenomenal consciousness to fend off potentially decisive objections to physicalism; Levine has convinced me, though, that they must do more.

In what follows I shall focus exclusively upon Levine’s central positive thesis, though there is much else in the book that is also worthy of discussion. According to this thesis, phenomenal consciousness exhibits two features that “both resist explanatory reduction to the physical: subjectivity and qualitative character” (p. 175). Let us start with qualitative character. Suppose that I am currently in the neural state which, as a brain scientist might put it, is the “neural basis” of red sensations in the sense that, whenever someone is in that state, he or she reliably occupies an internal state which seems to its owner to exhibit phenomenal redness. We can ask why my being in that neural state should mean that I am in a state which exhibits phenomenal redness. Why phenomenal redness and not phenomenal greenness or blueness—or just nothing at all? Nothing in the nature of the neural basis of red sensations seems to ensure that phenomenal redness must be its concomitant.

Now Levine is well aware that this demand for a reductive explanation of phenomenal redness in physical terms can be met. It is true that nothing in the nature of the neural basis of red sensations seems to ensure that phenomenal redness must be its concomitant; but that is just to say either that there is no *a priori* connection between the concept of the neural basis of red sensations and the concept of phenomenal redness, or that our current and no doubt incomplete grasp of the nature of the neural basis of red sensations and the nature of phenomenal redness reveals no necessary connection between them. But if, as the physicalist anticipates, we learn more of the nature of phenomenal redness, and if in particular we make the empirical discovery that enjoying phenomenal redness is one and the same thing as being in such-and-such neural state, or one and the same thing as being in such-and-such physically-realized functional state, then, given that I am in the right neural state, I cannot fail to enjoy phenomenal redness—just as my cup cannot fail to contain water given that in fact it contains H₂O.

So attention must be redirected to the alleged phenomenal/physical (or phenomenal/functional)

identities themselves: why is being in such-and-such physical (or functional) state identical with enjoying phenomenal redness, rather than with enjoying phenomenal greenness—or some state there is nothing it is like to be in at all? But this last question need not be answered, according to Levine, since he “generally endorse[s] the claim that pure identities are not suitable candidates for explanation” (p. 81). Nevertheless, he insists that there are such things as “gappy identities,” and that alleged phenomenal/physical (or phenomenal/functional) identity claims are examples. If an identity claim is gappy, then it “admits of an intelligible request for explanation” (p. 84).

But the very idea of a gappy identity is hard to understand. For since identities hold as a matter of metaphysical necessity, and since it only makes sense to ask why something is the case if it could have been otherwise, it makes no sense to ask why identities hold. And Levine offers no reason to suppose that this argument fails to apply in the case of phenomenal/physical (or phenomenal/functional) identities. So what does Levine mean when he claims that phenomenal/physical (or phenomenal/functional) identities give rise to an explanatory gap? The answer, I think, is that the kind of explanation he misses is a “how-possibly” explanation, rather than a “why” explanation; and that, in particular, he wants to know how the first-person concepts of phenomenal properties that we deploy in introspecting our own current phenomenal states could possibly turn out to co-refer with certain concepts drawn from physical science or functionalist psychology—as they must if any phenomenal/physical (or phenomenal/functional) identity is to be truly asserted. But a request for a “how-possibly” explanation must be made, I presume, within a context which indicates some reason for thinking that the phenomenon in question is not possible. So what is Levine’s reason for suspecting that first-person concepts of phenomenal states could not possibly co-refer with certain physical or functional concepts? Certainly it is not that our first-person concepts of phenomenal properties have descriptive—and richly descriptive—modes of presentation, so that in merely thinking of our own phenomenal states via such concepts we are thereby thinking of them as possessing certain features, features that arguably no physical or functional state could or does possess; for in Chapter 2 Levine is happy to deny that first-person concepts of phenomenal properties have descriptive modes of presentation at all.

To find Levine’s reason for suspecting that first-person concepts of phenomenal states could not possibly co-refer with physical or functional concepts, we must turn to his views about subjectivity, the second of the two features of phenomenal consciousness that he claims resist explanatory reduction to the physical. He writes:

... with phenomenal concepts, such as our concept of a reddish quale, there is a “thick,” substantive mode of presentation ... we have a fairly determinate conception of what it is for an experience to be reddish. This is ... a reflection of the subjectivity of conscious experience, the fact that my qualia are “for me” in a cognitively substantive and determinate way. (p. 84)

Note that when he speaks here of a “substantive” mode of presentation, he cannot mean a descriptive mode of presentation, since he disavows those in Chapter 2; later we will consider what he does mean. Note also that what he describes here is only one reflection of the “subjectivity of conscious experience”; there is a second.

Now clearly, according to Levine, the fact (if it is a fact) that “we have a fairly determinate conception of what it is for an experience to be reddish” is a puzzling phenomenon irrespective of whether it provides a reason for suspecting that first-person concepts of phenomenal states could not possibly co-refer with certain physical or functional concepts. However, he also thinks, I presume, that it does provide such a reason. But how? The answer is unclear. Part of it is that if phenomenal concepts have a “substantive” mode of presentation, then that undermines the physicalist suggestion that first-person phenomenal concepts are, or are rather like, demonstrative concepts; for in that case, intuitively, their modes of presentation would be non-substantial (p. 82). But another part of the answer seems to involve an appeal to introspection:

I am told that my concept of reddishness is really about a neurophysiological or functional property. I then wonder, as I ostend the reddishness of my visual experience, how could a functional or physiological state be *that*? (p. 83)

I must confess that I do not see what Levine is getting at here. Had these sentences come from a less hard-headed philosopher than Levine, I would construe them as expressing the view that introspection of a sensation (i) directly reveals its essence, and (ii) directly reveals that essence to be non-physical; but in several passages later in the book Levine explicitly rejects that view. Nor can these sentences be interpreted as alluding to some rich descriptive mode of presentation that first-person phenomenal concepts allegedly have, for the reason already given.

But let us return to the question of what Levine had in mind in the passage quoted above when he speaks of “a fairly determinate conception of what it is for an experience to be reddish” and claims that first-person phenomenal concepts have a “substantive” mode of presentation. The most explicit answer we get is perhaps this passage from the introduction:

When I think of what it is to be reddish, the **reddishness itself is somehow included in the thought; it's present to me.** This is what I mean by saying it has a “substantive” mode of presentation. (p. 8)



Let us suppose that it is true that first-person thoughts about reddishness somehow include the reddishness itself. Is this something that a physicalist view of the mind can accommodate? Let me suggest a way in which it could. Should this suggestion prove correct, then not only would (one aspect of) what Levine calls “subjectivity” turn out not to resist explanatory reduction to the physical, but nothing would be left of the claim that phenomenal character itself resists such reduction.

My suggestion is that, surprisingly, what Levine says about thoughts about reddishness might be literally true: reddishness might literally be a part of (some) introspective thoughts about reddishness. Begin by assuming that intentionalists about phenomenal properties (e.g. Tye, Lycan, Dretske) are correct to hold that reddishness (i.e. phenomenal redness) is a representational property—something like the property of non-conceptually representing that something external to oneself is red. On this view, then, sensations are representations of the external world or of one's body, and their phenomenal character is their representational content. We are then in a position to construe an introspective thought (of one import-

ant kind) about one's current red sensation as composed of two syntactic constituents: first, an operator which operates on bearers of propositional content, and whose meaning is something like “It appears to me now as if ...” or “I am now sensorily representing that ...”; and, second, a particular bearer of propositional content (namely, the red sensation itself) whose non-conceptual content is something like—though much more fine-grained than—“Something external to me is red.” This suggestion is not the familiar idea that introspective thoughts are, or are like, demonstrative mental sentences which demonstrate sensations, so that the propositions such sentences express are incomplete unless the sensations themselves are relevantly present. Rather, the suggestion is that introspective thoughts are not even complete mental sentences unless the sensations themselves are relevantly present, since the sensations are literally syntactic constituents of the introspective thoughts. It is as if I were to say “Tomorrow there is going to be a” and then hold up a photograph of a thunderstorm in progress; without holding up the photo (or doing anything else), my sentence would be incomplete. Notice that this suggestion is consistent with our also being able to report introspectively on our current sensations by tokening sentences that are completely conceptual, and whose content is something like “It appears to me now as if something external to me is some or other shade of red,” where, because the number of shades we can discriminate sensorily far exceeds the number of color concepts we possess, such a report always seems inadequate to its subject matter. Notice, also, that my suggestion could also explain how reddishness might literally be a part of a non-introspective thought about reddishness. If I think today about the reddish sensation I enjoyed yesterday, then we can suppose that in doing so I am today hosting some visual mental image that is reddish.

For Levine, however, the subjectivity of conscious experience has two aspects, and the very tentative suggestion of the previous paragraph addresses only the first. The second aspect of subjectivity is allegedly that “Qualia are such as to necessitate awareness of them” (p. 168). I am certainly inclined to agree that my phenomenal states are in some sense essentially “for me.” But physicalists might account for this feature in several ways. They might claim that it is an illusion, induced by the fact that my first-person thoughts

about my own current phenomenal states are to the effect that it appears to me that so-and-so, or that I myself am sensorily representing that so-and-so. Alternatively—or additionally—they might claim, with the intentionalist about phenomenal properties, that phenomenal states are representations, but then insist that their representational content is always in part about oneself (e.g. “Something external to me is red,” “Damage is occurring in my foot”).

Obviously I am not yet persuaded of Levine’s central positive thesis. But all philosophers of mind with an interest in consciousness, physicalist or not, will want to read this book so they can formulate their own response to his challenge (“physicalist or not” because of his provocative claim that “it isn’t really physicality that presents the problem”; p. 177). And they will want to read it, too, for the sake of the many shrewd and interesting discussions it contains of virtually all the canonical issues and positions familiar from the recent literature on phenomenal consciousness.

ANDREW MELNYK
University of Missouri
 USA

The subtlety of emotions

AARON BEN-ZE’EV
 Cambridge, MA: MIT Press, 2000
 ISBN 0-262-02463-2

Emotions are a hot topic in philosophy and the cognitive sciences right now. Several influential books on emotions have come out in the past few years, bringing increased attention and interest to this fascinating area of research. Many of these works have emphasized the *science* of emotion (for example, see Damasio, 1994, 1999; Griffiths, 1997; LeDoux, 1996), demonstrating the legitimacy of emotions as a subject of rigorous scientific inquiry. Until fairly recently many scientists eschewed the study of emotions, perhaps because they seemed too subjective to study with any scientific rigor or perhaps because they were thought to be tangential to work on important cognitive processes such as memory, attention and perception. It is now becoming clear (due, in part, to the work done by those cited above) that emotions influence processes such as memory and attention

and are themselves important features of cognition and behavior. Philosophers have a longer (albeit not well populated) history of thought on the emotions and the recent spate of philosophical books on emotions (for example, Elster, 1999; Goldie, 2000; Nussbaum, 2001, as well as the work under review) joins a line of philosophical thought on emotions stretching back to Aristotle and the Stoics.

Aaron Ben-Ze’ev’s *The subtlety of emotions* holds promise of integrating these sometimes quite different views of emotions. In the introduction to the book, Ben-Ze’ev states that “[f]urther progress in understanding emotions requires an interdisciplinary approach that combines a philosophical perspective with other types of scientific research (especially psychology), as well as with insights from other sources, such as folk wisdom and art” (pp. xiii–v). He references a wide variety of sources from philosophy, anthropology, sociology, psychology, and other areas (and liberally sprinkles the book with quotes from luminaries as varied as Thoreau, Yogi Berra, Theodore Roosevelt and Cindy Crawford). Ben-Ze’ev uses these different sources to present a detailed picture of the emotions in their social and personal contexts. He introduces the project of the book with this: “Emotions are highly complex and subtle phenomena whose explanation requires careful and systematic analysis of their multiple characteristics and components” (p. 3). Ben-Ze’ev presents a framework for characterizing and classifying the emotions which allows him to explain features and patterns associated with both the emotions in general and specific emotion types. The lengthy book is divided into two parts. Part I concerns the nature of emotions and addresses issues such as the rationality of emotions and the relationship between emotions and morality. Part II provides analyses (using the framework set out in Part I) of specific emotions such as envy, jealousy, pity, anger, hate and romantic love.

The analysis of emotions that Ben-Ze’ev offers in Part I is a variation on a cognitive-evaluative or appraisal theory of emotions in that it defines and individuates emotions in terms of the cognitive appraisals, evaluations or judgments involved. The approach borrows both from philosophical propositional attitude theories of Aristotle, William Lyons, Robert Solomon and others, and the appraisal theories of psychologists such as Richard Lazarus, Andrew Ortony and Keith Oatley (with whom one of the chapter sec-

tions is written). Ben-Ze'ev places special emphasis on the complexity and subtlety of emotions (hence the title of the book), and their sensitivity to social and personal contexts. Ben-Ze'ev argues that because of this, emotions in general as well as individual emotion types are best viewed as prototype categories. Category membership is determined by degree of similarity to a prototypical member or best exemplar with no element or elements distinguished as necessary or sufficient. Ben-Ze'ev describes a series of typical characteristics and components that define the category, emotion. His determinations of "typical" features are derived primarily from conceptual analysis, bolstered by some data from the social sciences. This same methodology is used for his analyses of individual emotions in Part II.

Ben-Ze'ev's conceptual framework highlights the roles social and personal relations play in the generation and unfolding of our emotions. He argues that emotions are responses to those changes in the environment which have been evaluated as significant in light of our personal concerns. Our personal concerns reflect our own values and perspectives, and typically involve our relationships with others. Ben-Ze'ev emphasizes the social, *comparative* nature of typical emotional concerns. We compare a present state either with a previous or future state, or with the state of another person, and in light of our perceived status or that of another. Emotions function to draw attention to such changes and marshal our response. Ben-Ze'ev describes emotions as typically short-term, intense transition states that express the narrow, personal, or "partial" perspective of the subject.

He argues that emotions possess elements of two basic mental dimensions: intentionality and feeling. The *intentionality* of emotions refers to their being directed toward or "about" some object (actual or imaginary) and consists of cognitive, evaluative and motivational components. *Feelings* are also an integral component of our emotion experiences; they have no intentional content, but instead express the subject's own state. These four components, cognition, evaluation, motivation and feeling, are each distinct yet complementary aspects of an emotion state. The *cognitive* component supplies information about the changes in our situation and a description of the emotional object. The information and description of the object is often distorted and reflects the "partial" nature of emotions. Emotions

are "partial" in the sense that our interpretation of the emotional situation and its object reflects a narrow, subjective perspective, as opposed to one that is impartial and broad. The *evaluative* component of emotion is of central importance because it is what distinguishes one emotion type from another. For example, guilt involves the evaluation of one's actions as having violated certain norms, while pride includes a positive evaluation related to self. The evaluative component evaluates or appraises the information presented by the cognitive component in light of our personal sets of values, goals and attitudes. Evaluation and cognition are each elements of what is experienced as a unified and instantaneous phenomenon. The way we evaluate a situation *motivates* us to respond in certain ways and determines what courses of action we take or are disposed to take.

In Part II Ben-Ze'ev applies the descriptive framework outlined in Part I to particular emotions, using it to both characterize and distinguish between similar but distinct emotion types such as envy and jealousy, and pity, compassion and mercy. For example, Ben-Ze'ev describes envy as a two-place relation between the subject and another whose attributes the subject lacks but desires to have. The central element of envy is the subject's perceived inferiority of self, where that inferiority in relation to the other's good fortune is evaluated as undeserved. Jealousy, on the other hand, is a three-part relation between a subject, the subject's mate and the mate's relationship with another. In jealousy the subject perceives that his unique relationship with his mate is being threatened. One way Ben-Ze'ev contrasts the two is by pointing out that envy is concerned with a *current* situation where we perceive our own inferiority in relation to another, whereas jealousy concerns a potential *future* loss. Therefore an envious person seeks to change the current situation, whereas a jealous person seeks to prevent change. Ben-Ze'ev's descriptions of the different emotion types in Part II is thorough, although occasionally somewhat taxonomic in tone, and he explores some interesting related issues. For example, he discusses the moral values of different emotions, pointing out, for instance, that envy is often viewed as extremely morally negative whereas jealousy, rooted as it is in loving relationships to others, is viewed as less so.

The subtlety of emotions is an informative, wide-ranging discussion of the nature of emotions. However, Ben-Ze'ev's decision not to critically

engage the emotions literature (p. xv) may frustrate some readers. He bypasses some central debates and issues, and his discussion of others can be disappointingly brief. This is due in part to the descriptive focus of the project; he is concerned primarily with “making plausible generalizations” about emotions (p. 5), and is content to leave it at that. The broad descriptive project is largely successful; Ben-Ze’ev draws out interesting relations and patterns that provide further insight into the characteristics of emotions. However, this may not be satisfying to those readers who want something beyond plausible generalizations and are interested in a more in-depth, rigorous examination of the issues.

For example, in laying out the distinctions between emotions, moods, sentiments and affective disorders (Chapter 4), Ben-Ze’ev describes moods and affective disorders as possessing “primitive” or “general” intentionality in that “the intentional object is diffuse and difficult to specify” (p. 87). Given his discussion of intentionality earlier, it is not clear what he means by “primitive” intentionality or the sense in which he thinks moods have intentional objects. The intentionality of moods is a subject of debate in the affect literature and has presented problems for some theories of affect that make intentional states necessary components of all affective states (for discussion see Armon-Jones, 1992; Griffiths, 1997; Lormand, 1985; Sizer, 2000). As Armon-Jones puts it, “an affective state which lacks a differential, selective focus on a determinate particular, can no more be described as having an object, than a light which indiscriminately illuminates everything in the room can be described as a spotlight” (Armon-Jones, 1992, p. 133). Therefore it is rather frustrating that Ben-Ze’ev does not say more to clarify his views on these issues.

Appraisal theories of emotion have encountered criticism from psychologists and neuroscientists who argue that many emotions unfold independently of higher cognitive processing—the sorts of processing thought to be required for cognitive appraisal or evaluation. This is seen as a significant challenge to appraisal theories and has produced a great deal of debate in the literature (see especially Lazarus, 1982, 1984; Zajonc, 1980, 1984).

The criticism (made by LeDoux, 1996, Zajonc, 1980, 1984; among others) is that emotions can occur independently of or prior to the involve-

ment of the sorts of higher cognitive processes associated with beliefs, desires and judgment formation. Joseph LeDoux has argued that the brain contains several different pathways along which emotion-relevant information travels. The “quick n’ dirty” route proceeds from the thalamus directly to the amygdala, triggering emotion responses *prior* to that information reaching higher cognitive processing centers. This research seems to present a direct challenge to emotion theorists—like Ben-Ze’ev—who argue that cognition and evaluation are defining components of *all* emotions.

Ben-Ze’ev briefly addresses this criticism by distinguishing between constitutive and causal appraisal theories (p. 72). His is a constitutive view. He states that in typical cases evaluations both cause and constitute the emotion, but admits that some non-typical emotions are “generated by merely having the suitable facial or physiological features” (p. 72). A constitutive appraisal theory may avoid the criticism, but the issue begs for more discussion than Ben-Ze’ev provides. In particular one would like to see him clarify how such non-typical emotions could be integrated with the rest of his analysis. There is another puzzle about the nature and role of evaluations in Ben-Ze’ev’s account. He distinguishes between *deliberative evaluations* which are relatively slow, involve conscious processes under voluntary control and operate on verbally accessible, semantic information, and *schematic evaluations* which are fast, automatic, below conscious awareness and use a tacit, elementary evaluative system (pp. 57–58). Similar distinctions are made by Damasio, LeDoux, Ekman and others. Ben-Ze’ev holds that emotional evaluations are *schematic*. In other words, they are nondeliberative and spontaneous, the products of automatic appraisal mechanisms (p. 59), and do not require higher cognitive processing. While Ben-Ze’ev does not seem to pursue this route, this distinction could give him another way to respond to the criticism that emotions can occur independently of higher cognitive processing.

However, Ben-Ze’ev’s claim that emotion evaluations are schematic merely prompts further questions. It is not clear how Ben-Ze’ev would reconcile the complex sets of cognitions and evaluations central to his characterization of emotions, and sensitive to myriad personal and social variables, with the idea of an evaluation system that operates in a fast, rough-and-ready fashion, with no higher cognitive processing involved. Ben-

Ze'ev does state that the evaluation system is "based upon readymade structures or schemes of appraisal which have already been set during evolution and personal development ... Since the evaluative patterns are part of our psychological constitution, we do not need to create them; we just need the right circumstances to activate them" (p. 58). This is an intriguing claim, but one that merits further justification and discussion. Ben-Ze'ev devoted an entire chapter to convincing us that emotions are incredibly complex and sensitive to context, so it is hard to see how "tacit, elementary" evaluation systems can pick out, record and respond to these many different richly nuanced evaluative patterns. This is a difficult issue to resolve. However, the fact that Ben-Ze'ev does not say more on this topic may leave the reader feeling that there is a profound disconnect between his characterization of the subtlety of emotions and the nature of the mechanisms involved.

Readers looking for a thorough discussion of certain emotion debates and issues may be disappointed by *The subtlety of emotions*. However, the book does an impressive job of drawing together ideas from many different fields in the service of presenting a detailed characterization of emotions and the complex social and personal factors that structure our emotional interactions. It shines a light on some emotions that are rarely discussed in the literature and makes important observations about emotions and the relationships between different emotion types.

References

- ARMON-JONES, C. (1992). Affect, objects and rationality. *Journal for the Theory of Social Behavior*, 22, 129–143.
- DAMASIO, A. (1994). *Descartes' error: emotion, reason and the human brain*. New York: Avon Books.
- DAMASIO, A. (1999). *The feeling of what happens*. New York: Harcourt.
- ELSTER, J. (1999). *Alchemies of the mind: rationality and the emotions*. New York: Cambridge University Press.
- GOLDIE, P. (2000). *The emotions: a philosophical exploration*. London: Oxford University Press.
- GRIFFITHS, P. (1997). *What emotions really are*. Chicago: University of Chicago Press.
- LAZARUS, R. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, 37, 1019–1024.

- LAZARUS, R. (1984). On the primacy of cognition. *American Psychologist*, 39, 124–129.
- LEDoux, J. (1996). *The emotional brain*. New York: Simon & Schuster.
- LORMAND, E. (1985). Toward a theory of moods. *Philosophical Studies*, 47, 385–407.
- NUSSBAUM, M. (2001). *Upheavals of thought: the intelligence of emotions*. New York: Cambridge University Press.
- SIZER, L. (2000). Towards a computational theory of moods. *British Journal for the Philosophy of Science*, 51, 743–769.
- ZAJONC, R. (1980). Feeling and thinking: preferences need no inferences. *American Psychologist*, 35, 151–175.
- ZAJONC, R. (1984). On the primacy of affect. *American Psychologist*, 39, 117–123.

LAURA SIZER
Hampshire College
Amherst, MA 01002–5001
USA

The paradox of sleep: the story of dreaming

MICHEL JOUVET
Cambridge MA: MIT Press, 1999

The French physiologist Michel Jouvet is one of the big names of the "golden age" of physiological sleep and dream research which flourished after the 1953 discovery of REM sleep and its close association with dreaming. His book *The paradox of sleep* offers an interesting personal view on the developments and changing ideas in the field, written in an accessible style. The original French edition *Le sommeil et le rêve* was published in 1993, the English translation in 1999. There has been rapid progress in the field since then and therefore the book is no longer quite up to date in all respects. The recent brain imaging experiments on REM sleep are hardly even mentioned, for example.

Two main themes run through the book: first, the history of ideas and discoveries in sleep and dream research, and second, the quest for understanding the function of dream sleep.

Changing views on sleep and dreaming

The explanation of the relationship between

dreaming and sleep originated with the dualistic idea that the body goes through a periodic death—the ancient myth of Hypnos (sleep) being the brother of Thanatos (death). But the soul is ever active and not dependent on the state of the material body. Dreaming was taken to be an expression of the immaterial soul in continuous motion. In the 19th century Alfred Maury overthrew the ancient dualistic conception, claiming that dreaming is an episodic state, intermediate between sleep and wakefulness. However, it was not until the 1950s that physiological research instruments (such as polysomnography and single-cell recordings) capable of testing the hypothesis became available. Aserinsky and Kleitman, who in 1953 discovered the REM stage of sleep, were obviously still influenced by Maury's conception, and therefore interpreted stage REM as the return of "light" sleep, a state of half-waking and half-sleeping.

Jouvet was not satisfied with this conception, but, based on his own data, suggested in 1959 that REM sleep was neither sleeping nor waking but a *third* state of the brain, and a very curious one at that. It was a state of relatively deep sleep during which the brain is highly active, the voluntary muscles are paralyzed, and rapid eye movements and the subjective experiences of dreaming frequently occur. But it is not just a light form of ordinary slow-wave sleep. Thus, Jouvet introduced the concept of *paradoxical sleep* to denote this curious third state. The concept is undoubtedly one of his main contributions to the field.

The mechanisms of dreaming

Jouvet observes that the study of the mechanism and function of dreaming can be conducted either "from the inside" at the level of the subjective content of dreams or "from the outside" by objective experimental approaches. However, his book falls short of presenting an adequate overview of content analysis studies of dreams (see e.g. Domhoff, 1996; Strauch & Meier, 1996). That is a pity, for it is difficult to see how the mysteries of dreaming could be solved without a systematic and detailed description of the actual *explanandum*, subjective dream experience. Instead, Jouvet mostly concentrates on the underlying physiological mechanisms. The spirit of the times in the 1960s, according to him, was: "Dreaming had thus become a physiological phenomenon and we thought that neurophysiological methods would

enable us to resolve the mystery of its mechanisms" (p. 159).

As Jouvet himself admits now, this approach has not led to an understanding of what dreams are or why they exist. I suspect that the neglect of dream phenomenology is one reason for the lack of progress: the level of organization at which dreaming proper occurs had been overlooked, as if the phenomenon of dreaming could be described and explained at the purely physiological level.

In accordance with the idea that a physiological approach might explain dreaming Jouvet accepts the hypothesis that the mechanisms of dreaming are the same as the mechanisms of REM sleep. This is a controversial assumption. Its validity largely depends on the *definition* of dreaming: what actually counts as a dream? Depending on the definition, the correlation between physiological stage REM and subjective dreaming appears to be either stronger or weaker. In sleep and dream research these issues continue to be hotly debated. It seems that some kind of graded scale classifying subjective experiences during sleep with regard to their complexity and organization might turn out to be useful to clear up the problem of the definition of a dream (Nielsen, 2000; Revonsuo, 2000a). Furthermore, a better understanding is required also of the physiological criteria for REM sleep. In a recent paper, Nielsen (2000) suggests that dream experiences during NREM sleep may coincide with "covert" REM sleep: short episodes of sleep which show some but not all of the standard criteria for REM sleep.

The fact is that we still do not know which neurophysiological phenomena are necessary and sufficient for the presence of subjective experiences during sleep, and which methods could be used to detect signals from those phenomena. Indeed, this question comes very close to the more general question of uncovering the neural correlates of consciousness in the brain (Metzinger, 2000), a hot topic in current cognitive neuroscience. It seems that dream research could play a role in that effort (for recent ideas along these lines, see Schwartz & Maquet, 2002).

A window to the dream world

A key question regarding the mechanisms of dreaming is whether the REMs and the PGO-wave activity characteristic of paradoxical sleep are random, chaotic forms of activation or expressions of an organized code. Jouvet's fascinating studies

on the oneiric behavior of cats during REM sleep have thrown light on this. If the normal muscular atonia accompanying REM sleep is removed (by destroying the inhibitory nucleus in the brain stem that normally brings atonia about), a most surprising thing happens: when REM sleep commences, the cat opens its eyes and raises its head, and then engages in unpredictable but stereotyped behavioral patterns such as stalking, predatory aggression, fear and rage responses and so on. However, the cat is not aware of its real surroundings—it is blind and deaf to it—but appears to interact with an entirely imaginary environment and its invisible objects. The oneiric behaviors resemble play, but the behavioral patterns are internally determined.

Is oneiric behavior a window to the animal's subjective dream world? In any case, such an interpretation of the data is extremely tempting. It has gained increasing support since 1986, when a corresponding syndrome, called REM sleep behavior disorder (RBD), was identified in humans. Patients suffering from RBD have a neurological deficit which weakens motor inhibition during REM and consequently the patients forcefully act out their dreams and nightmares, completely unaware of the physical surroundings. They typically wake up after having collided with something and hurt themselves. They usually recall a vivid dream, and the actions carried out within the dream perfectly correspond to the actions carried out by their physical bodies just before waking up. RBD in humans and oneiric behaviors in the cat provide strong support for the hypothesis that paradoxical sleep involves the activation of structured, integrated motor programs whose effects normally remain trapped within the central nervous system.

The function of dreaming

According to Jouvet, the problem with dreaming is that it is a phenomenon without a function, at least, without a known function of its own. What's worse, the prospects for understanding its functions are not good, for "it is difficult to understand how dreaming could provide an evolutionary advantage when it corresponds to a state when the animal is most vulnerable" (p. 111). Jouvet remarks that there are as many theories of dreaming as there are theorists, and that every research school in the 1960s tried to find a function for dreaming. He reviews some of these theories and then presents his own bold hypothesis. He admits

that it does not entirely solve the mystery of dreaming "and will doubtless soon seem just as erroneous as all the others that repose in the graveyard of dream theories" (p. 133). In a nutshell, Jouvet's hypothesis is that dreaming (or rather the physiology of REM sleep) is the repeated genetic reprogramming of our individuality. During dreaming a repetitive program (expressed as PGO-waves and REMs) on the one hand wipes out certain aspects of what we have learned and on the other reinforces features which are compatible with the genetic program. This would preserve the behavioral variation in individuals who have been exposed to the same environment. In animals, paradoxical sleep could contribute to the programming of species-specific instinctive behavior, such as observed during oneiric behaviors. Jouvet presents evidence that the patterns of REMs and oneiric behaviors are similar in genetically identical or similar individuals, but entirely different in genetically different individuals. Thus, they might express a genetically controlled code for reprogramming the neural networks in the brain.

One of Jouvet's main arguments for the theory is the fact that neurons in the adult mammalian brain do not divide anymore but have only one fate: death. Thus, the genetic program cannot be expressed in the adult brain through continuous neurogenesis; the genetically programmed organization of the nervous system must be maintained through some other means. Unfortunately the doctrine about the absolute impossibility of neurogenesis in the adult brain seems to have become outdated very recently, for new evidence has been presented that even in the mammalian brain, neuronal stem cells do exist and they can produce new neurons throughout life. It would be interesting to hear Jouvet's own evaluation of this new evidence: how serious a blow against the genetic reprogramming theory of REM sleep does this revolutionary new finding constitute? Unfortunately during the writing of the book he must have been unaware that any such findings were in the offing.

Jouvet's dream theory has one serious weakness: it gives no functional role to the subjective experience of dreaming, the actual phenomenal content of dreams. In the theory, dream images are seen as mere epiphenomena that happen to be elicited during the genetic reprogramming. Thus, one may wonder whether this is a genuine theory of the function of *dreaming* at all, or only of the

function of REM sleep. I fear that Jouvett fails to take the phenomenal content of dreams seriously enough. Furthermore, he fails to draw a proper distinction between REM sleep as a physiological phenomenon and dream experience as a subjective mental phenomenon.

Contrary to Jouvett, I believe that the mystery of the function of dreaming might be much better illuminated if the subjective content of dreams is taken into account and treated as a biological level of organization of its own (I have called it “the phenomenal level”), and further if the “function” of dreaming is regarded as “the original biological function” of dreaming (i.e. the feature of dreaming that had adaptive value for ancestral humans in their environment). In that case the function of dreaming can only be discovered by finding out about the organization and dynamics at the phenomenal level (the content of dreams) and about the environment our ancestors were living in and that their dreams must have been about.

I have recently suggested that such an analysis leads to the view that dreaming is an organized nocturnal simulation of the individual’s world, predisposed to simulate events in the environment that threaten the individual’s reproductive success in some way (Revonsuo, 2000b). This tendency to simulate threats is reflected, for example, in the predominance of negative over positive elements in dreams, the universality of nightmares, the regular appearance of repetitive post-traumatic nightmares after life-threatening events during wakefulness, and universal threat themes in dreams (chase, attack, falling from a high place, losing valuables, being late, and other greater and lesser concerns we regularly have dreams about). My hypothesis is that during evolution this sort of a system was selected for because the individuals who were equipped with it received automatic and perfectly safe mental training in threat-perception and threat-avoidance skills. Individuals whose threat coping skills thus became more efficient were more likely to leave offspring. We are descended from those ancestors whose brains used the idle time during sleep to prepare themselves to survive the worst threats encountered in their environment. The theory shows how the subjective content of dreams may have become highly relevant to ancestral reproductive success, a requirement that any convincing theory of the biological function of dreaming must fulfill.

Although Jouvett’s own theory of the function of dreaming has not gained wide support, his

fascinating observations about oneiric behaviors in the cat can be neatly accommodated by the threat simulation theory. We only need to assume that members of different species automatically rehearse the species-specific survival skills that are most critical to the reproductive success of each species. Stalking, predatory aggression, etc. observed in the dreaming cat surely fits the bill. Human RBD patients often act out dreams in which they forcefully attempt to escape from some mortal threat, thus manifesting threat-avoidance behaviors typical for (ancestral) humans in life-threatening situations.

Sleep and dream theories: what are the correct levels of explanation?

Jouvett appears to be aware of the importance of finding the correct level of description at which a phenomenon is to be explained. He criticizes the direction in which sleep research has been going: basic physiological research such as he himself has done is no longer in vogue; instead, molecular approaches reign. In his opinion, one mistake between 1960 and 1980 was to assume that research into micro-level mechanisms and structures would automatically lead to the function of dreaming and sleep. One example of this trend was the monoaminergic theory of sleep that was launched in 1969. For a while everything in sleep research was explained by referring to the neurotransmitters that had just been discovered. But when the story at that level became more and more complicated as more and more different neurotransmitters were found, the hope for a simple localization of function was lost. In the last paragraph of the book, Jouvett summarizes the insight he has gained:

It is obvious that the understanding of the logic of sleeping and dreaming depends on finding the appropriate level. It does not seem that the sleeping or dreaming brain will be explained by putting together our knowledge about all the molecules involved, or of a genetic program. Complexity has its own laws. (p. 178)

This seems to echo the message from the philosophy of neuroscience and mechanistic biological explanation: a crucial step in explaining any biological phenomenon is to find the correct levels of description and explanation. Jouvett seems

to admit that his generation of sleep researchers did not quite succeed in discovering all the relevant levels involved. I would suggest that the first reasonable step in the search for the appropriate levels of description is to take the phenomenal level of subjective experience as seriously as all the other biological levels of organization in the brain.

The *Paradox of sleep* is an enjoyable read for anyone interested in sleep and dreams. It is written in an accessible, almost light style. Jouvett's colorful opinions and witty remarks of the world of science create the impression that this is an honest personal account of a hard-working researcher's life-work, with its successes and frustrations. The book is not even intended as a representative or balanced review of the whole field, but this is one man's personal story of dreaming, a story well worth relating.

Acknowledgement

Antti Revonsuo is supported by the Academy of Finland (project #45704).

References

DOMHOFF, G.W. (1996). *Finding meaning in dreams: a quantitative approach*. New York: Plenum.

METZINGER, T. (Ed.) (2000). *Neural correlates of consciousness*. Cambridge, MA: MIT Press.

NIELSEN, T.A. (2000). A review of mentation in REM and NREM sleep: "covert" REM sleep as a possible reconciliation of two opposing models. *Behavioral and Brain Sciences*, 23, 793–1121.

REVONSUO, A. (2000a). The contents of consciousness during sleep: some theoretical problems. *Behavioral and Brain Sciences*, 23, 995–996.

REVONSUO, A. (2000b). The reinterpretation of dreams: an evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences*, 23, 877–901.

SCHWARTZ, S. & MAQUET, P. (2002). Sleep imaging and the neuropsychological assessment of dreams. *Trends in Cognitive Sciences*, 6, 23–30.

STRAUCH, I. & MEIER, B. (1996). *In search of dreams: results of experimental dream research*. New York: SUNY Press.

ANTTI REVONSUO

Department of Philosophy

Center for Cognitive Neuroscience

University of Turku

FIN 20014 Turku

Finland

Copyright of Philosophical Psychology is the property of Carfax Publishing Company and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.