j.apa

Mark Solms

REVISION OF DRIVE THEORY

Several deep revisions of Freud's theory of the drives are proposed: (1) Drives are conscious and are in fact the source of all consciousness. (2) Drive energy is equated with variational free energy and is therefore quantifiable in principle. (3) There are not two drives but many, seven of which may be described as "emotional" as opposed to "bodily" drives. (4) All drives are self-preservative or preservative of the species; there is no death drive at work in the mind. This means, at the mechanistic level, that all drives are homeostatic and anti-entropic. (5) The great task of mental development is to supplement instinctual predictions about how our multiple drive demands may be met and reconciled with each other. This work is done by learning from experience, mainly through voluntary behavior, which is governed by conscious feelings.

Keywords: affect, drive theory, Freudian theory, id, instincts, memory, neuropsychoanalysis, neuroscience, pleasure principle, repetition compulsion, free energy, homeostasis, self-organizing systems

Psychoanalytic drive theory is in need of fundamental revision. Many analysts today have not only lost confidence in Freud's classical conception of the drives (1915a, 1920) but have abandoned drive theory altogether, thereby conflating *Freudian* drive theory with drive theory per se. The price we pay for this is too great; rather than abandon drive theory, we must revise it.

The raison d'être of psychoanalysis was in Freud's view to bring the phenomena of mental life within the purview of science; to treat the mind as just another part of nature. His pioneering efforts to that end were predicated on the assumption that we human beings are, after all, a species of

Director of Neuropsychology, University of Cape Town and Groote Schuur Hospital; Co-Chair, International Neuropsychoanalysis Society; Chair of Research, International Psychoanalytical Association; Science Director, American Psychoanalytic Association.

The author thanks Gyuri Fodor for helpful comments on an earlier draft of this paper. Submitted for publication August 26, 2020.

DOI: 10.1177/00030651211057041

animal. In this sense, Freud is rightly described as a "biologist of the mind" (Sulloway 1979).

If the mind is part of nature, then it is *embodied* (what else can it be?); and it is from this simple premise that Freudian drive theory starts. Human mental functions, like those of all other species, ultimately serve the evolutionary imperatives of survival and reproductive success. This approach enabled Freud to see the obvious, shocking though it may have been to his Victorian contemporaries: the sexual drive plays the same fundamental role in human psychology that it does in animal behavior. To abandon drive theory is to abandon this insight, which is the essential connection between psychoanalysis and the body, and, indeed, with the life sciences as a whole.

Freud's classical definition of drive (1915a) is so well known that it hardly needs quoting[1]:

> "Drive" appears to us as a concept on the frontier between the mental and the somatic, as the psychical representative of the stimuli originating from within the organism and reaching the mind, as a measure of the demand made upon the mind for work in consequence of its connection with the body [1915a, pp. 121–122].

All the introductory points I have just made about the embodied nature of mind are illustrated in this definition, but it reveals something else, too. The phrase "a measure of the demand made upon the mind for work" is also a *psychophysical* statement, which alludes to Freud's first ambition, namely to "furnish a psychology that shall be a natural science: that is, to represent psychical processes as quantitatively determinate states of specifiable material particles, thus making those processes perspicuous and free from contradiction" (Freud 1895, p. 295). Although Freud gave up his attempt to represent mental processes as *material* ones, he never abandoned the deeper scientific goals on which this endeavor was based. Those goals are encapsulated in another quotation that again is so well known that it barely needs repeating. I am referring to the goals of the "Helmholtz school of medicine," of which Freud's revered teacher Ernst

[1]In line with the editorial policy followed in the *Revised Standard Edition,* I am throughout this paper replacing Strachey's translation of *Trieb* as "instinct" with "drive." However, my quotations are referenced to the pagination of the original *Standard Edition,* since the revised edition is still in press.

von Brücke was a founding member.[2] The mission of the Helmholtz school was described as follows by Emil du Bois–Reymond, another founding member, in 1842:

> Brücke and I pledged a solemn oath to put into effect this truth: "No other forces than the common physical and chemical ones are active within the organism. In those cases which cannot currently be explained by these forces one has either to find the specific way or form of their action by means of the physical-mathematical method or to assume new forces equal in dignity to the chemical-physical forces inherent in matter, reducible to the forces of attraction and repulsion" [Du Bois–Reymond 1842, p. 108].

This goal became the impetus behind Freud's fundamental theory of psychoanalysis, his "metapsychology," which, as he explained at the time, was intended to replace metaphysics (i.e., to rescue psychology from philosophy and claim it for science).[3] This is important, because Freud's metapsychology has suffered much the same fate as his drive theory. That is, many psychoanalysts today, by abandoning *Freudian* metapsychology, have abandoned *any* attempt to cast the structure and functions of the mind in lawful, mechanistic terms (cf. Holt 1967; Gill and Holzman 1976; Klein 1976; Schafer 1976). Again, in my view, the price we pay for this is too great. There can be little doubt that Freud would have deemed it a tragedy.

The aspect of Freud's metapsychology that has been most roundly rejected is precisely the part that overlaps with his drive theory, namely, what he called the "economic point of view." This is the *quantitative* viewpoint mentioned both in the opening sentence of the 1895 Project, quoted above, and in Du Bois–Reymond's solemn oath. It is alluded to also in Freud's definition of drive as *a measure* of the demand for mental work. Although, as I have said, Freud abandoned "*for the present*" (1915b, p. 175; emphasis in original) his attempt to translate psychodynamics into neurophysiology (see Freud 1920, p. 60), he never abandoned his deeper

[2]In his *Autobiographical Study,* Freud (1925) wrote: "In Ernst Brücke's physiological laboratory, I found rest and full satisfaction—and men, too, whom I could respect and take as my models: the great Brücke himself, and his assistants" (p. 10). In a postscript to "The Question of Lay Analysis," Freud (1927) added that Brücke "carried more weight with me than anyone else in my whole life" (p. 253).

[3]When Freud first introduced this term, he clarified that it refers to a level of explanation that incorporates both psychological and biological facts (letter to Fliess. March 10, 1898 [Masson 1985]).

assumption that the mind, like everything else in nature, must be driven by an *energy* of some kind. Thus, he clung to his notion of "drive energy"; a *psychical* energy that was capable of increase, diminution, displacement, and discharge, and therefore possessed all the characteristics of a quantity, "though we have no means of measuring it" (Freud 1894, p. 60). This is the most direct example of Freud's commitment to the ideals of the Helmholtz school: in cases that cannot yet be explained in terms of physical and chemical forces, one has "either to find the specific way or form of their action by means of the physical-mathematical method *or to assume new forces equal in dignity.*"

I have gained the impression, over decades of work in this field, not only that most psychoanalysts have rejected the notion of "drive energy" and the associated "economic point of view" but that many of us are positively embarrassed by these concepts. Let me be clear: what is at stake here is not only whether psychoanalytic theory is compatible with the basic tenets of physical science (whereby nothing can perform work without using an energy to do so), but also whether we are unwittingly consigning ourselves to mind/body dualism (whereby the mind exists somehow independently of the physical universe).[4] A far better approach, in my view, is to do what Freud expected his intellectual descendants to do, which is what we do in every other science: revise and correct his basic concepts in line with unfolding evidence. That is my purpose here.

Since this paper carries over into psychoanalysis assumptions derived from other branches of knowledge, I must briefly, before proceeding, address the often expressed objection that it is unpsychoanalytic or even antipsychoanalytic to do so (see, e.g., Blass and Carmeli 2007). When it comes to drive theory, this objection is more easily dismissed than in other such matters, for the reason that Freud always acknowledged in the

[4]For Freud (1915a, pp. 121–122), as quoted above, drive was a concept "on the frontier between the mental and the somatic"; it was the foundation stone of his monism. This is not the place to enter into lengthy philosophical discussions; however, to avoid misunderstanding, let me clarify that metapsychological constructs (like "drive") are *functional* entities that manifest in both psychological and physiological *phenomena* (e.g., feeling states and their correlated neurochemical events). Like Freud, I take the position of dual-aspect monism: psychological events are no more caused by physiological ones than thunder is caused by lightning. Our goal, therefore, is not to reduce psychological phenomena to physiological ones, but rather to explain their *common underlying mechanism* in natural-scientific terms. This requires us to transcend the languages of psychology and physiology both, and to replace them with the deeper (nonphenomenal) abstractions of statistical physics. As Galileo said: "The book of Nature is written in the language of mathematics."

case of the "frontier" concept of drive that we are *obliged* to rely on knowledge from other fields:

> I am altogether doubtful whether any decisive pointers for the differentiation and classification of the drives can be arrived at on the basis of working over the psychological material. This working-over seems rather itself to call for the application to the material of definite assumptions concerning the life of the drives, and it would be a desirable thing if those assumptions could be taken from some other branch of knowledge and carried over to psychology [Freud 1915a, p. 124].

Many other such statements could be cited. It is abundantly clear from Freud's writings on this topic that he drew heavily on the biology and psychophysics of his day. That is why he felt so uncertain about his conclusions—"it should be made quite clear that the uncertainty of our speculation has been greatly increased by the necessity for borrowing from the science of biology"—and why he readily conceded that future biological findings "may be of a kind that will blow away the whole of our artificial structure of hypotheses" (Freud 1920, p. 60). In short, in the case of drive theory, there can be no doubt that we may, indeed must, carry over into psychoanalysis findings derived from other branches of knowledge.

I. A BRIEF REVIEW OF CLASSICAL DRIVE THEORY

Freud's drive theory, well known to psychoanalysts, does not require detailed exegesis here. I will briefly outline only the major issues that seem to need revision: (1) the fundamental mechanism of drive; (2) the bodily sources of drive; (3) classification of the drives; (4) the relationship between drive and pleasure-unpleasure; and (5) the relationship between drive and conscious qualities in general. Since these points are discussed so frequently in Freud's writings, I am not going to clutter my summary with detailed bibliographic references, though I will employ quotation marks wherever I use Freud's own words. These are words he used many times; I will make referenced quotations only in relation to some specific and less familiar issues.

(1) The Fundamental Mechanism of Drive

Freud believed, from the first to the last of his writings, that the fundamental mechanism of drive was to *discharge* as rapidly as possible the

excitations that impinge upon the mind from the interior of the body.[5] The purpose of such discharge was to keep the level of the relentless somatic sources of excitation as low as possible. Freud linked this discharge mechanism of drive with the psychophysical notion of a "freely mobile" energy—the driving force of his "primary process" of mental functioning. His abiding model for this mechanism was that of the reflex arc. In his earliest writings he formulated it as a "principle of neuronal inertia," that is, as the reflexive tendency of nerve cells to divest themselves of excitation. In his last writings, he formulated it in terms of the "Nirvana principle," that is, the striving of the nervous system to rid itself of excitation as quickly as possible and thereby return to the quiescent state—and ultimately to an inert, inorganic state. In a word, for Freud, drives were *entropic*. That is, drives (in his first [1915a] theory, but the "death drive" only in his second [1920] theory; see below)[6] were governed by the Second Law of Thermodynamics, according to which the energy in every system naturally *dissipates*. In this connection, it is evident that Freud's distinction between "free" and "bound" states of drive energy built directly upon the physics of Helmholtz, who introduced the term "bound energy" into thermodynamics and indeed played a major role in the formulation of its First Law.[7] It is also important to note in this connection that Freud's "constancy principle," in terms of which the mental apparatus tolerates a tonic store of energy, which is kept as constant (and low) as possible, and which becomes its supply of "bound energy" (and thereby of the secondary process), is a *compromise* between the primary process of mental functioning and the "exigencies of life" (which later became the "reality principle"). Freud's constancy principle was not a manifestation of the

[5]A final word on philosophical premises: "the mind" is merely a *subjective observational perspective* upon "the body"; it is the being of the body. The scientific question arising from this premise is: why does it feel like something to be a human body but not to be, say, a computer? In other words, what does sentient subjectivity *do* and how does it arise? The answer to this question, as we shall see, has everything to do with the biological function of feeling. Please note: my account of the mechanisms through which feeling comes about is not intended to be *metaphorical.* These are the actual, causal mechanisms of sentient subjectivity. (For the fuller treatment that this immensely complex issue requires, see Solms [2021]. My topic here is Freudian drive theory, not the mechanistic origins of selfhood.)

[6]In Freud's second drive theory, the fundamental entropic tendency of drive is represented in the life drives, too, by way of their "conservative" nature; that is, through their compulsive tendency to revert to a *previous state* of things (previous to the impingement of an excitation). See the "compulsion to repeat."

[7]"In my opinion this distinction [between free and bound drive energy] represents the deepest insight we have gained up to the present into the nature of nervous energy, and I do not see how we can avoid making it" (Freud 1915b, p. 188).

basic tendency of drives themselves, which remained an unrelenting tendency toward discharge.

(2) The Bodily Sources of Drive

Freud distinguished between the "pressure," "aim," "object," and "source" of a drive. The "pressure" referred to its *quantitative* aspect, to the fact that it was a "measure" admitting of degrees. The "aim" referred to its basic mechanism, discussed above, which was to discharge, that is, to get rid of the demand. The human infant, like all animals, is equipped with reflexes and instincts[8] that serve this purpose, but since these innate mechanisms cannot by themselves satisfy the exigencies of life, at least not in every context, they must be supplemented by learning—through "experiences of satisfaction." This applies all the more to the "object" of a drive, which, as Freud pointed out, is its most variable aspect; the object of a drive is almost always specified through *learning*.

The "source" of every drive, even of the death drive, is the "cellular processes of the body." (Freud equated the life and death drives with anabolism and catabolism, respectively.) In regard to the life drives and their sources, he placed special emphasis on mechanical stimulation of the skin and other sensory organs, and in particular of the mucous membranes. This applied not only to the libidinal drives but also to the self-preservative ones, in his first drive theory. For example, he imagined that irritation of the lining of the stomach and dryness of the pharynx gave rise to the drives of hunger and thirst, respectively. Such stimulation, according to Freud, resulted in the release of "chemical products,"[9] including specifically sexual ones. He took the view (which he called "a provisional assumption that we cannot escape in the theory of the drives") that "excitations of two kinds arise from the somatic organs, based upon differences of a chemical nature. One of these kinds of excitation we describe as being specifically sexual, and we speak of the organ concerned as the 'erotogenic zone' of the sexual component drive arising from it" (Freud 1905, p. 168).

[8]The fact that instincts *satisfy* drive demands demonstrates just how unfortunate it was that Strachey translated *Trieb* as "instinct." The German word for instinct is *Instinkt* (see Solms 2018b).

[9]Interestingly, Freud predicted on this basis that it might become possible in the future (our present) to exert a "direct influence" on the drive economy by means of psychopharmaceuticals. In this connection, Freud likened mental illnesses to "toxic processes," and he emphasised the similarities between the neuroses and Graves' disease (hyperthyroidism) and Bright's disease (acute or chronic nephritis). Significantly, in this respect he also drew attention to the libidinal intoxication that is caused by ingestion of alkaloids like cocaine, something about which he knew much from personal experience.

Thus, for example, sucking at the breast simultaneously satisfies the self-preservative drive of hunger and (the oral component of) the libidinal drive "anaclitically." That is, the libidinal function *leans upon* the nutritional function.[10] The same applies to defecation, in the case of the anal component drive, which provides libidinal satisfaction while simultaneously meeting a self-preservative need. According to Freud, it is only in the process of development, and therefore through learning, that the component libidinal drives detach themselves—and then only partially—from these self-preservative functions (later called "narcissistic"). In this way, said Freud, the libidinal component drives mature and coalesce to achieve "genital" sexuality, in service of the biological function of reproduction. This developmental ideal is, however, not always attained, in which case the component drives are said to be "fixated" on pregenital aims and objects.

(3) Classification of the Drives

Freud freely admitted that he found his classification of the drives to be problematic. It was especially in this respect that he thought "it would be a desirable thing if those assumptions could be taken from some other branch of knowledge and carried over to psychology." However, in the absence of definite guidance from the biology of his time, he felt obliged to formulate some *provisional* hypotheses.

For some reason (perhaps the injunction to "assume new forces equal in dignity to the chemical-physical forces inherent in matter, *reducible to the forces of attraction and repulsion*"), Freud considered it necessary to classify the drives on a *dualistic* basis.[11] Clinical experience demanded a classification that could accommodate the ubiquity of *conflict* in mental life, but it seems obvious that this requirement does not necessarily imply a two-way conflict (e.g., one can be sexually aroused by an object, and annoyed by it, and dependent upon it (i.e., "attached" to it), all at the same time).[12] Be that as it may, Freud initially drew a dichotomy between the "self-preservative" and "libidinal" drives, and he linked this with the

[10]*Anlehnung,* which Strachey translated as "anaclisis," literally means "leaning-on-ness" (see Solms 2018b).

[11]"The analogy of our two basic drives extends from the sphere of living things to the pair of opposing forces—attraction and repulsion—which rule in the inorganic world" (Freud 1940, p. 149).

[12]Later we shall learn that these inclinations are manifestations of three different drives.

contrasting needs of the individual and the species. However, with his development of the concept of narcissism (self-love), he recognized that self-preservation was itself libidinally driven, and he therefore felt obliged to combine the libido with self-preservation under the rubric of what he now called the "life drives" (or "Eros"). This led him to postulate a new and deeper principle "beyond the pleasure principle" (see section [4] below), which he called the "Nirvana principle." This principle was the direct expression of entropy, that is, of the natural tendency to dissipate, which all living things must oppose if they are to maintain their existence. This fundamental entropic tendency was counteracted by the life drives, the aim of which was "to establish ever greater unities and to preserve them."

The conflicting demands of the constructive and destructive drives were said to operate throughout the organism, through processes of "fusion" and "de-fusion." Thus, for example, self-preservative aggression (and indeed, motility in general) was conceptualized by Freud not as a primary disposition but rather as the product of a fusion between libido and the death drive, which is thereby diverted away from the subject and toward objects. The function of the self-preservative drives in general thereby became a deflection of *external* threats away from the organism: "to assure that the organism shall follow its own path to death, and to ward off any possible ways of returning to inorganic existence other than those which are immanent in the organism itself" (Freud 1920, p. 39).

> In biological functions the two basic drives operate against each other or combine with each other. Thus, the act of eating is a destruction of the object with the final aim of incorporating it, and the sexual act is an act of aggression with the purpose of the most intimate union. This concurrent and mutually opposing action of the two basic drives gives rise to the whole variegation of the phenomena of life [Freud 1940, p. 149].

On this basis, Freud reconceptualized the development of the sexual function—from oral "cannibalism," through anal-sadistic control and phallic superiority, to mature genital love—as the product of progressively more successful fusions between destructiveness and libido (i.e., progressive taming of our ambivalence toward objects). Falling ill was now conceptualized as a *regression* of this progress—that is, as drive de-fusion.

The extent to which Freud was obliged to speculate about the physiology of the drives is evident from the fact that he thought epileptic seizures might be a clinical expression of drive de-fusion. The same applied to his speculations about "germ plasm," which were formulated before the discovery of DNA. The following quotation illustrates the limitations of the biological knowledge then available:

> The id, guided by the pleasure principle—that is, by the perception of unpleasure—fends off [the life drives] in various ways. It does so in the first place by complying as swiftly as possible with the demands of the non-desexualized libido—by striving for the satisfaction of the directly sexual trends. But it does so in a far more comprehensive fashion in relation to one particular form of satisfaction in which all component demands converge—by discharge of the sexual substances, which are saturated vehicles, so to speak, of the erotic tensions.[13] The ejection of the sexual substances in the sexual act corresponds in a sense to the separation of soma and germ-plasm. This accounts for the likeness of the condition that follows complete sexual satisfaction to dying, and for the fact that death coincides with the act of copulation in some lower animals. These creatures die in the act of reproduction because, after Eros has been eliminated through the process of satisfaction, the death drive has a free hand for accomplishing its purposes [Freud 1923, p. 47].

(4) The Relationship between Drive and Pleasure-Unpleasure

This is trenched upon by the opening sentence of the above quotation. As is well known, Freud initially postulated a direct psychophysical ratio between increasing and decreasing drive tension on the one hand and unpleasure and pleasure on the other. (Incidentally, it is curiously unclear whether this applied only to the libidinal drive. It seems that Freud conceptualized the affective valencing of self-preservative needs in terms of "anaclisis"; see section [2] above. However, the question fell away when self-preservation was understood as an expression of self-love.) This conception of drive pressure in relation to pleasure-unpleasure makes intuitive sense, and it accordingly became the foundation of the "pleasure principle," in terms of which animals like us are attracted to pleasurable aims and objects and repelled by unpleasurable ones. (Notice, once more, the attraction-repulsion polarity.) It will become important

[13]See section (2) above: Freud (1905, p. 213) imagined that the "the accumulation of the sexual substances creates and maintains sexual tension; the pressure of these products upon the walls of the vesicles containing them might be supposed to act as a stimulus upon a spinal centre, the condition of which would then be perceived by higher centres and would give rise in consciousness of the familiar sensation of tension."

later to notice that Freud saw the pleasure-unpleasure series as a *continuum:* the organism is compelled to seek ever more pleasure. It is only due to the constraining influence of the "reality principle" that this compulsive urge is tamed, as the ego learns to tolerate a constant store of libidinal energy, which characterizes the constancy principle and gives rise to the secondary process.

Freud soon was beset by doubts concerning this directly proportional ratio between drive pressure and affective valence. His canonical counterexample was the problem of foreplay (which he explained by the theory of pregenital modes of satisfaction), but, more generally, it was the problem that sexual excitement itself, presumably characterized by *increasing* libidinal pressure, is typically experienced as *pleasurable*. Freud was never able to resolve this fundamental issue; he could ultimately do no better than to suggest that perhaps pleasure and unpleasure correspond not to the absolute level of drive tension but rather to fluctuations in it over a given period of time. (He equated this with amplitude.) The following quotation from *Beyond the Pleasure Principle* illustrates the extent of the problem, and it also introduces another one:

> Here might be the starting point for fresh investigations. Our consciousness communicates to us feelings from within not only of pleasure and unpleasure but also a peculiar tension which in its turn can be either pleasurable or unpleasurable. Should the difference between these feelings enable us to distinguish between bound and unbound processes of energy? or is the feeling of tension to be related to the absolute magnitude of the cathexis *within a given unit of time?* Another striking fact is that the life drives have so much more contact with our internal perception—emerging as breakers of the peace and constantly producing tensions whose release is felt as pleasure—while the death drives seem to do their work unobtrusively. The pleasure principle seems actually to serve the death drives [Freud 1920, p. 63.]

The new problem emerges with the closing sentence. In the Project, Freud had initially equated the pleasure principle with neuronal inertia; its relationship with entropy became even more clear when he equated it, as he did here, with the Nirvana principle. But since the Nirvana principle was served by the death drive, this had the dismaying implication that pleasure was associated with satisfaction not of the life drives but of the death drive. That is why Freud ultimately felt obliged, as we saw in the previous section, to construe the pleasure associated with orgasm—a phenomenon so closely connected with reproduction (from the biological

standpoint) that it surely serves the purposes of *life*—as a deathly event. This is a clear indication that Freud might have taken a wrong turn, somewhere, in developing his theory of the drives.

Before moving to the relationship between drive and conscious qualities in general, there is one last issue concerning its relationship with pleasure-unpleasure that must be addressed. Affect, for Freud, is not reducible to "oscillations in the tension of drive needs" alone; it is more complex than that. Here is his fullest definition of it:

> What is an affect in the dynamic sense? It is in any case something highly composite. An affect includes in the first place particular motor innervations or discharges and secondly certain feelings; the latter are of two kinds—perceptions of the motor actions that have occurred and the direct feelings of pleasure and unpleasure which, as we say, give the affect its keynote. But I do not think that with this enumeration we have arrived at the essence of an affect. We seem to see deeper in the case of some affects and to recognize that the core which holds the combination we have described together is the repetition of some particular significant experience. This experience could only be a very early impression of a very general nature, placed in the prehistory not of the individual but of the species (Freud 1916–1917, p. 395).

So affect, for Freud, included an additional dimension, which he conceptualized under the heading "phylogenetic memory." A pivotal example is "castration anxiety," which he saw as an inherited disposition based not solely (and sometimes not at all) in individual experience but rather in prehistoric events: the repeated experience by our ancestors, over many generations, of an alpha male in a primal horde who castrates or tries to castrate his sons when they reach puberty and threaten his preeminent sexual position. According to Freud, the descendants of these sons inherit unconscious memories of those traumatic events, which take the form of what he called "primal fantasies," which form the nucleus of the system *Ucs*. It is precisely in relation to these inherited dispositions that Freud used the word "instinct" (*Instinkt*) rather than "drive" (*Trieb*). Having drawn attention to this aspect of affect, I will discuss it no further, precisely because it is *not* the expression of a drive. I will defer discussion of this matter to a companion paper titled "Revision of the Theory of the Biological Origins of the Oedipus Complex" (Solms in press).

(5) The Relationship between Drive and Conscious Qualities in General

Although Freud (1894) first introduced a quantitative factor into mental life with reference to "quotas of affect" that cathect memory traces

of ideas, just one year later (1895) he had already adopted a different view: drive energies—now called "*Q*η"—were in the first instance *unconscious;* they become conscious only when they reach the "upper story" of the mental apparatus, the cortical system "ω" (his later system *Cs*.). Freud held to this view for the remainder of his scientific life; not only to the view that drive energies were unconscious in themselves but also that consciousness was an exclusive property of the *cortex*:

> The process of something becoming conscious is above all linked with the *perceptions* which our sense organs receive from the external world. From the topographical point of view, therefore, it is a phenomenon which takes place in the *outermost cortex of the ego*. It is true that we also receive information from the inside of the body—the feelings, which actually exercise a more peremptory influence on our mental life than external perceptions; moreover, in certain circumstances the sense organs themselves transmit feelings, sensations of pain, in addition to the perceptions specific to them. Since, however, these sensations (as we call them in contrast to conscious perceptions) also emanate from the terminal organs and since *we regard all these as prolongations or offshoots of the cortical layer,* we are still able to maintain the assertion made above. The only distinction would be that, as regards the terminal organs of sensation and feeling, the body itself would take the place of the external world [Freud 1940, pp. 161–162; emphasis added].

The distinction between unconscious drive energies and conscious perceptions of them must not be conflated with Freud's notion of drive as "a concept on the frontier between the mental and the somatic." For Freud, drive energies operating *within the mind* (in the id and in the unconscious parts of the ego and the superego) were unconscious in all their vicissitudes until they reached the superficial system *Cs*. Crucially, this was how he made sense of the fact that drive demands emanating from one source can become "confluent" with those from another source; so that, for example, sexual longing can be transformed into hunger. This also explains how unconscious drive demands can be misconstrued by consciousness, through processes like condensation, displacement, reversal of affect, and the like, as occur in the dreamwork and in defense.

The same theoretical claim, namely, that drive energies are unconscious in themselves, explains why Freud insisted that affects are *necessarily conscious* phenomena. There is no such thing as an "unconscious emotion." This is because affects, for Freud, were perceptions of the process whereby drives are *discharged,* so that a drive remains unconscious

until it is discharged, at which point it is consciously perceived and is no longer a drive but an affect.[14]

Freud saw consciousness, in both its manifestations—endogenous affect and exogenous representation—as a *perceptual* process, whereby stimuli impinging upon a terminal organ (whether it be visceral or sensory) are propagated in quantitative form to the cerebral cortex, which is where they are registered as "qualities." (Accordingly, from 1917 onward, Freud combined the interoceptive system *Cs*. with the exteroceptive *Pcpt*. into a single system *Pcpt.-Cs*., which perceives affects on its internal surface and objects on its external one.) Where, then, do the *differences* in our conscious qualities—both affective and perceptual—spring from? Freud's answer was:

> Everything points to the sense organs, whose qualities seem to be represented precisely by different periods [i.e., amplitudes] of neuronal motion. The sense organs act not only as *Q*-screens, like all nerve-ending apparatuses, but also as *sieves;* for they allow the stimulus through from only certain processes with a particular period. [Freud 1895, p. 310].

The same applied to the interoceptive terminal organs. Freud's view was:

> Are we to suppose that the different drives which originate in the body and operate on the mind are also to be distinguished by different *qualities,* and that that is why they behave in qualitatively different ways in mental life? This supposition does not appear to be justified; we are much more likely to find the simpler assumption sufficient—that the drives are all qualitatively alike and owe the effect they make only to the amount of excitation they carry, or, perhaps, in addition, to certain functions of that quantity. What distinguishes from one another the mental effects produced by the various drives may be traced to the difference in their *sources* [Freud 1915a, p. 123; second emphasis added].

I trust I am not alone in finding this confusing. How can the quantitative aspect of different drives (their voltages, as it were) coalesce during the passage from the terminal organs through the "lower stories" of the mind to the system *Pcpt.-Cs.,* while their differential qualitative aspects (their "periods" or amplitudes) remain distinct?

[14]See Freud (1915b): "Affects and emotions correspond to processes of discharge, the final manifestation of which are perceived as feelings. In the present state of our knowledge of affects and feelings we cannot express this difference more clearly" (p. 177).

A bigger problem, however, was this:

> The id, cut off from the external world, has a world of perception of its own. It detects with extraordinary acuteness certain changes in its interior, especially oscillations in the tension of its drive needs, and these changes become conscious as feelings in the pleasure–unpleasure series. It is hard to say, to be sure, by what means and with the help of what sensory terminal organs these perceptions come about. But it is an established fact that self-perceptions—coenaesthetic feelings and feelings of pleasure–unpleasure—govern the passage of events in the id with despotic force. The id obeys the inexorable pleasure principle [Freud 1940, p. 198].

If affects are by definition conscious (see above), then how can it be true (1) that the passage of events in the id obeys the inexorable pleasure principle *and* (2) that the id and its drives are unconscious? What is the point of a pleasure principle "in" the id if it does not feel it? And if the drives are felt as pleasure and unpleasure only when they reach the perceptual superficies of the ego, then the sole way for the pleasure principle to govern the passage of events in the id is through top-down control. But that is the role of the reality principle. Surely the pleasure principle involves bottom-up influence? Once again we are on the brink of theoretical incoherence—a good sign that we might have taken a wrong turn.

II. PROPOSED REVISIONS

One faces a dilemma when drawing on knowledge from a discipline one's readers are not familiar with. Too much technical detail can make the exegesis overwhelming, but too little can render it dogmatic. This applies especially to contested issues, which exist in every science, and which are bound to be highly specialized issues for the very reason that they represent the cutting edge of a science. The solution I adopt here is to rely on the conclusions (and technical terminology) of the most widely known and respected authorities, while acknowledging that not *everyone* agrees with them and that current knowledge is by no means complete. That said, few would disagree that the leading neuroscientific theorists on the topic of drive, in the current state of the field, are Jaak Panksepp (1998), Donald Pfaff (1999, 2005), Antonio Damasio (2010, 2018), and Karl Friston (2009, 2013). I will draw also on the work of Bjorn Merker (2007), though he is less well known. These five scientists represent diverse

methodologies, ranging from deep brain stimulation to computational modeling; they also represent both clinical and experimental methods, in both humans and other animals. As in the previous section, I am not going to clutter my exegesis with a plethora of references; the grounding references just provided will have to suffice. Since I cannot possibly do justice to the complexity of the issues within the constraints of a journal article, readers wishing to master the underlying science can consult those sources. (Alternatively, for a book-length overview of the technical issues, see Solms 2021.) I am painfully aware, however, that the account I am about to provide might well sound dogmatic, due to its highly condensed nature.

I am using letters to designate the sections in this part, rather than numbers, as I did in Part I, because the topics do not coincide exactly. It is not possible to discuss our current understanding of the fundamental mechanism of drive, for example, without simultaneously considering its relationship to pleasure-unpleasure. I will cover the following topics: (A) the fundamental mechanism of drive and its relationship to pleasure-unpleasure; (B) the relationship between drive and entropy; (C) classification of the drives and the bodily sources of drive; (D) the relationship between drive and conscious qualities in general.

(A) The Fundamental Mechanism of Drive and Its Relationship to Pleasure-Unpleasure

The fundamental mechanism of drive is *homeostasis*. This concept was introduced into biology by Walter Cannon (1926), but it became widely known only in the 1930s, long after Freud had formulated his basic ideas. Cannon's concept should not be confused with Claude Bernard's notion of a *milieu intérieur,* which had more in common with Freud's "constancy principle" than with his understanding of the fundamental mechanism of drive. Had Freud been familiar with the concept of homeostasis, I believe, he would have formulated his drive theory differently, because, at a stroke, it solves many of the problems that bedeviled him.

Homeostasis maintains the system that it regulates within "expected" bounds. Thus, just as your domestic thermostat drives an engine that maintains your desired room temperature, so too the temperature control system of your brain works to maintain core body temperature within its expected range of 97.7–99.5 degrees Fahrenheit. A crucial difference between these two examples, though, is that the expected bounds of your body are also its *viable* bounds. In other words, they are crucial for the

survival of your body as a system; unlike your home, when your body overshoots or undershoots its expected temperature range by too large a measure, it ceases to exist as a system. That is, it dies.

Deviations from the expected ranges function physiologically as "error signals" and represent *demands for work*. However, not all such demands can be equated with "drive," not only for the reason just mentioned (viz., that drives are underwritten by a biological scale of values—that it is *good* for the system to survive and reproduce—while domestic thermostats are not) but also because many such error signals, even in biological systems, do not represent demands upon *the mind* for work. To be clear: many homeostatic systems exist (even in the body) that are regulated *autonomically*. There is nothing mental about them. Blood-pressure regulation is a clinically notorious example: you (your mind) knows nothing of your body's blood-pressure undershoots and overshoots until it is too late. Accordingly, unless you monitor your blood pressure artificially (externally), you cannot become aware of the error signals and therefore can do nothing to correct them. Similarly, your body obtains energy by metabolizing glucose from adipose tissues according to need, but no drive is involved in that process. The demand upon your mind to find further supplies in the outside world, by contrast, *is* a drive—in exactly the sense that Freud (1915a) defined it—in this case, a drive called "hunger." To differentiate these two types of demand for work, the nonmental and the mental, I will use the terms "need" and "drive" respectively.

What characterizes the difference between a need and a drive (which lie on opposite sides of what Freud described as the "frontier" between the somatic and the mental) is well conceptualized by a term derived from statistical physics, happily a relatively simple one: "uncertainty." Blood-pressure error signals always evoke the same stereotyped response in the body (changes in heart rate and vasodilation or vasoconstriction). This is because the need to correct aberrant blood pressure is met monotonously, by reliable and certain measures, making it a *predictable* business. The mental work demanded by drive is required only when such stereotyped algorithms (which predict adequate solutions to error signals with certainty) are not available. That is when uncertainty prevails.

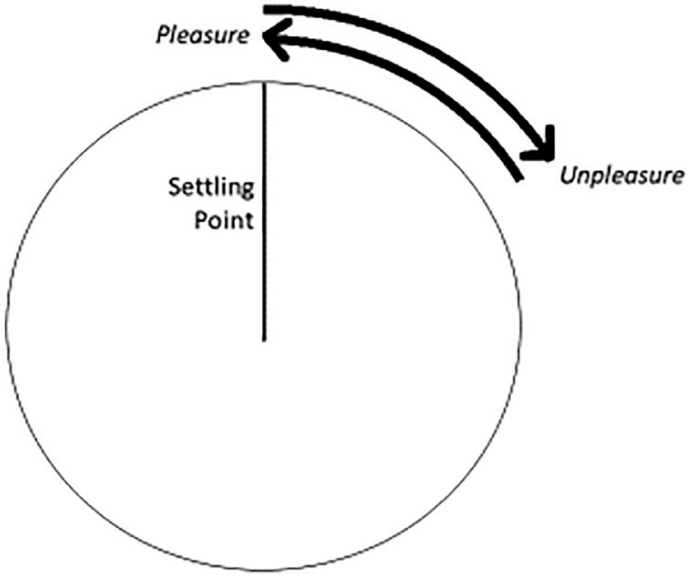
Here is an example to illustrate what the crossing of the somatic/mental frontier entails. Respiratory control is normally autonomic: so long as the levels of oxygen and carbon dioxide in your blood stay within expected viable bounds, you don't have to be aware of your need to

breathe. When blood gases exceed these normal limits, however, when autonomic mechanisms cannot cope with an emergency situation, respiratory control forcibly intrudes upon consciousness in the form of an acute feeling called "air hunger." This is a drive.[15] Unexpected blood gas values are an indication that *volitional action* is required. It is urgently necessary to remove an airway obstruction from your throat or to get out of a carbon-dioxide-filled room. At this point, the need to breathe enters consciousness, via an inner warning system that you experience as *alarm*—specifically, in this case, "suffocation alarm." Why does it enter consciousness? The answer is: because there is no response available that predicts with certainty that it will achieve the necessary outcome. The adequate solution remains *uncertain*. In other words, *choices* must be made.

Choices (the defining feature of *voluntary* as opposed to automatic control) must be grounded in a value system, which determines whether a selected action is "good" or "bad." (If there were no such value system, the behaviors would be random.) Our biological value system, as mentioned before, is that *survival is good and death is bad*. This seems to be the biological purpose of feeling: it tells you whether your actions are good or bad within this scale of values. The outcomes of actions that enhance your chances of survival and reproductive success feel good (this is "pleasure") and their opposites feel bad (this is "unpleasure"). Feelings register the consequences of your choices, as you make them, here and now. This enables you to determine, in the heat of the moment, whether one action is better or worse than another. In the example of air hunger, the regulation of your blood gases becomes conscious when you don't possess an automatized solution to maintain your physiologically viable bounds. In your rush to escape from a carbon-dioxide-filled room, for instance, how do you know where to turn? You have never been in this situation before (in *any* burning building, let alone this specific one), so you cannot possibly predict what to do. Now you must decide whether to go this way or that, up or down, etc. You make such choices by *feeling your way through the problem:* the feeling of suffocation waxes or wanes,

[15]The transition from glucose metabolism to hunger, just described, entails a similar transition. Core body temperature control provides another example: as your temperature rises, your body (not you) down-regulates it by perspiring and panting, but beyond a certain limit you *feel* too hot; then you must *do* something, like leave the kitchen.

Figure 1. Feeling homeostasis: The "settling point" represents the organism's viable bounds



depending on whether you are going the right way or not—that is, depending on whether the availability of oxygen increases or decreases.

So it seems that what distinguishes autonomic need from mental drive is the necessity for feeling. This answers the fundamental question posed earlier: what does sentient subjectivity *do?* Unpleasure tells the organism that things are getting worse (i.e., drive pressure is increasing), while pleasure tells it that things are getting better (drive pressure is decreasing), *in unpredicted situations,* and thereby underwrites the enormous adaptive advantages that a capacity for voluntary behavior bestows upon an organism.

This restores the simple psychophysical relationship between drive and pleasure-unpleasure that Freud thought he had to abandon. But please note: if the basic mechanism of drive is homeostasis, then the *ideal* state of the organism is not pleasure but rather *satiation*. "Satiation" is reached when viable bounds are regained; then the feeling at issue is resolved and it disappears from the radar of consciousness. This entire mechanism is illustrated in Figure 1.

Crucially, the mechanism I have just described drives not only voluntary behavior but also *learning from experience*. Learning is governed by

the Law of Affect,[16] which states that "if a behavior is consistently accompanied by pleasure it will increase, and if it is consistently accompanied by unpleasure it will decrease." For example, if you are a firefighter who has been in many burning buildings, you will gradually become less uncertain about what to do (you will have recourse to better predictions) and, accordingly, you will be less overwhelmed by affect (you will have less need of it, since affect is *felt uncertainty*). This applies quite generally, of course, and not only to suffocation alarm.

Learning, on this view, is a matter of *prediction*. We do not learn just because we can; we learn because we must. We learn how to meet our needs in the world. Memories are about the past but they are *for* the future. They predict the future on the basis of past experience. I will return to this important topic—which introduces the fundamental role of mental "representation"—in later sections. First, though, I must draw attention to the implications for Freudian drive theory of the simple mechanism I have just described.

The first implication is that *drives are conscious,* by definition. To me, it makes no sense to speak of hunger or thirst or suffocation alarm if you do not feel them. If you do not feel hungry, you are not hungry. It is true that the bodily *needs* that ultimately give rise to hunger and thirst (energy balance, sodium/water balance) are in the first instance monitored and regulated unconsciously, but these are not the kinds of processes that Freud denoted by the term "the unconscious." They remain on the bodily side of his frontier between the mental and the somatic. I propose that we use the term "nonconscious" for such nonmental processes.

I likewise propose that we define "affect" as the *subjective* aspect of "drive." I know this is a radical departure, but I do not see how we can avoid it. The distinction between the two terms nevertheless remains significant. In my view, the subjective aspect of the functioning of the mental instrument defines the observational perspective on it that is adopted by psychoanalysis; it defines the limits (and methods) of our discipline. The very same things that we psychoanalysts study from the subjective viewpoint, including unconscious intentionality, can be studied from the objective viewpoint, as neurobiologists do.[17] The requirement to study

[16]This is a mischievous rewording of the behaviorist Law of Effect, which speaks of "rewards" and "punishments" but does not define what they are because it denies the existence of feelings (see Panksepp and Biven 2012).

[17]The inter-discipline of neuropsychoanalysis aims to integrate the subjective and objective perspectives.

them objectively *and* subjectively is illustrated by such fundamental facts as the following.

I have said that the basic value system of biology is that it is good to survive and reproduce and bad not to do so. This explains *objectively* why pleasurable feelings came to be attached, through natural selection, to sexual behaviors—and to orgasm in particular (see above). However, these are not the values that motivate us *subjectively* to engage in such behaviors. When we have sex, we are not trying to perform our biological duty; on the contrary, we are frequently trying *not* to reproduce while having sex. This is because what motivates us subjectively are *feelings,* not the evolutionarily conserved mechanism that explains, objectively, why we have such feelings. This is not a semantic point. Confusion on this score gave rise to the mistaken belief, by Freud and a century of his followers, that *psychological* sexual maturity coincides with reproductive (heterosexual) "genitality." This is plainly not true, as surely any psychoanalyst can attest: it is perfectly possible to have psychologically mature homosexual relationships. The same erroneous assumption—namely, that biological causes must be psychologically represented, albeit unconsciously—gave rise to the mistaken belief that castration anxiety (as all the "primal fantasies") arises from unconscious *phylogenetic memories*. There are no such things as inherited memories, a neurobiological impossibility; but, as I have said, I will defer that topic to another paper (see Solms in press).

The second implication of the basic mechanism of drive is that what we are striving for is *satiation,* not *pleasure*. I will explain the details in section (C), when we reconsider the classification of the drives; for now I am making only the general point: pleasure predicts satiation; it tells you that you are heading in the right *direction,* but it is not the *end* of a drive. The pleasure-unpleasure series is not a continuum but an oscillator (see Figure 1).

Freud correctly inferred that the aim of every drive is satisfaction, but he incorrectly (I believe) equated satisfaction with pleasure. This had many baneful consequences. Not least of them was his assumption—since it is clear that many compulsive human behaviors are aimed not at pleasure but rather at what might be termed quiescence—that there must be a *deeper* principle in mental life, "beyond the pleasure principle." This became his Nirvana principle. I hope it is clear from what I have said,

first, that the pleasure principle is an extension of the Nirvana principle, that it *serves* the Nirvana principle. Second, I hope it is clear that, far from being deathly, satiation or "nothingness" (i.e., no needs making demands on the mind for work) is the ideal state of all self-preservative and libidinal creatures: it represents the *successful* outcome of work performed in response to drive demands.[18] Third, therefore, there is no need to invoke the existence of a separate "death" drive that serves the Nirvana principle; it is served by the "life" drives *and it represents their ideal state*. To be clear, there is no need (or justification) to conclude that "*the aim of all life is death*" (Freud 1920, p. 38; his emphasis). Here, again, Freud was conflating psychology with biology. The individual organism does not aim for death, psychologically, although it is our inevitable biological destination.

This is not to say that the psychological phenomena that Freud associated with the death drive do not exist. They appear to be of two kinds: developmental and pathological. The "repetition compulsion" (and the developmental phenomena associated with it, which Freud drew attention to, such as children wanting to repeat endlessly the same games and stories) seems to me simply an expression of the compulsion to *reduce uncertainty* (i.e., to increase predictability) mentioned above—and to be elaborated further in sections (B) and (D) below—which lies at the heart of "learning from experience." The *pathological* phenomena that Freud drew attention to are self-destructive behaviors of various kinds: the phenomena of malignant narcissism, negative therapeutic reactions, addictions, suicidality, and other such omnipotent mental formations. For the sake of brevity, since this is not a clinical paper, I will make a swift generalization. What these phenomena have in common (apart from the fact that they are not the expressions of a natural drive, precisely because they are pathological) is the fact that they are attempts to achieve satiation by "short circuit"; that is, they are attempts to satisfy the demands made upon the mind for work *without actually doing the work*. In other words, they are attempts to evade the reality principle, which is indeed a dangerous (and potentially fatal) thing to do. These are failures of *ego* functioning.

[18]Many readers will object at this point that satiation is not the goal of all motivated behavior. This is partly because the principle applies only in the long run (i.e., satiation is, on average, the goal of motivated behavior) but more so because I have not yet had opportunity to introduce the multiple drives that conflict with each other, and that include a default-mode SEEKING drive, the homeostatic settling point of which is proactive engagement with uncertainty (e.g., with novelty).

I have space for just one brief example. In section (C) I will introduce an "attachment" drive called PANIC/GRIEF, the command neuromodulator of which is beta endorphin acting on mu opioid receptors. Like all drives, PANIC/GRIEF makes demands upon the mind for work. The work in this case is to establish reunion with a missing caregiver. As we know, this can be difficult, especially for children who grow up in less than ideal circumstances. One way of dealing with the demand (which is felt as separation distress) is to *avoid* the real-life difficulty by replacing the desired object with an opiate drug (e.g., morphine or heroin). In this way you can achieve satiation (psychologically: the removal of separation distress; physiologically: mu opioid receptor binding) not by *real* reunion with the lost object, through mental work, but rather by short-circuiting the reality principle; i.e., by trying to trick or cheat it, which always ends badly. (E.g., heroin does not really take care of you.) This is deathly, but it is not the expression of a drive; it is a mechanism of *defense*.

(B) The Relationship between Drive and Entropy

I have said that the function of homeostasis is to maintain an organism within its viable bounds. Deviations from these bounds are entropic, and the mental work demanded by drives must return the animal to them. Hence, the work is *anti*-entropic. Moreover, it must be *effective* work.

"Entropy" is expressed in different ways in different disciplines. In thermodynamics it is described in terms of temperature and heat exchanges between a system and its environment, and in information science it is described in terms of information exchanges. Information exchange entails a system asking questions of its environment and receiving answers from it. (The aim of this is to reduce uncertainty; a system needs to ask questions only when it cannot predict an outcome with certainty.) The *fundamental* definition of entropy in modern physics, therefore, revolves around probability. Probability and predictability are two sides of the same coin. This applies both to thermodynamic exchanges and to information exchanges (see Jaynes 1957). The classical example is that of a compressed gas from a small cylinder being introduced into a larger chamber. As the gas disperses in the chamber—which it does naturally—entropy increases. Stated in terms of probabilities, this is because, as the molecules spread out to occupy all the possible positions in the larger container, it becomes increasingly difficult to predict the location of each and every molecule. The probability that a particular molecule will

occupy a particular location becomes increasingly unpredictable. As the locations become increasingly unpredictable, more *information* is required to describe the microstates of the system. As a result of this, the overall state of the system has become more uncertain. Information is quantified in "bits" (binary digits, usually denoted by "1" and "0," i.e., "yes" and "no"). This means, in the classical example, that more yes/no questions must be asked and answered to determine the positions of the molecules in the chamber; *more information must flow* before the position of each molecule can be specified.

To apply this concept from statistical physics to the biology of drive, if the function of a drive is to counteract entropy in a biological system, then it must act to minimize the information flow required to maintain its survival (i.e., it must reduce the number of questions that need to be asked by the system). Stated differently, the drive pushes the system to *minimize its uncertainty,* to render environmental events *more predictable*. This is not complicated; it just means that the drive must counteract the system's entropic tendency to disperse itself across all its possible states. Contrary to this tendency, the drive must restore and maintain the system's *expected* states.

To achieve this, as we know, the system must perform work, and this work must be effective. The energy in a system that is *usefully* employed (i.e., employed in effective work) is called "bound" energy, and the energy that is *not* usefully employed is called "free" energy. This, too, should sound familiar to you (see section [1] above). Like entropy, "free energy" is defined differently in different disciplines, but it always means the same thing, fundamentally: the energy in a system that is not currently employed in *effective* work is "free." What this means is that if the system is going to survive as a system (i.e., if it is going to avoid dissipation), then *the free energy in the system must be bound*. This is the "free energy principle."

In thermodynamics we speak of Helmholtz free energy, in chemistry we speak of Gibbs free energy, and in information science we speak of Friston free energy, also known as "variational" free energy. When Friston applied his free energy principle to the nervous system, he caused a revolution, the full effects of which have not yet been felt, at least not by nonspecialists, but it has resulted in his having become (objectively) the most influential neuroscientist in the world.[19] The revolution I

[19]The influence of scientists (the number of times their work is cited in the specialist literature) is measured by their "h-index." At the time of writing, Friston's index is 236. To put this into perspective, Einstein's is 110, Darwin's is 119, and Newton's is 62. Psychoanalysts will be pleased to know that Freud's is 286; so, if we reclassify Freud as a *neuro*psychoanalyst, he (rather than Friston) would be the most influential neuroscientist in the world today!

am referring to is the insight that the entire nervous system functions according to a single principle, the free energy principle, derived from statistical mechanics. This is important for psychoanalysts because the laws of probability apply equally to physiological *and* psychological events. Somatic and mental events are therefore explicable in terms of a single unifying principle. What the free energy principle achieves, therefore, is nothing less than to render mental life *quantifiable,* at least in principle.[20] The ramifications of this are enormous. For example, in my opinion, it brings the infamous "hard problem" of consciousness into the purview of statistical physics (see Solms 2021). (The hard problem concerns the question of how conscious states can be caused by and cause physical states.)

What concerns us here, though, is that Friston himself has equated his conception of free energy with Freud's "drive energy," in terms of which drive energy is isomorphic with "expected free energy" (see Carhart-Harris and Friston 2010). In other words, Friston's free energy principle achieves Freud's greatest ambition, namely, "to furnish a psychology that shall be a natural science: that is, to represent psychical processes as *quantitatively determinate* states of specifiable material particles, thus making those processes perspicuous and free from contradiction" (Freud 1895, p. 295). You will recall that Freud thought an energy was at work in the mind that is capable of increase, diminution, displacement, and discharge, and therefore possesses all the characteristics of a quantity, "though we have no means of measuring it" (Freud 1894, p. 60). Now we can measure it (for the grounding equations, see Solms and Friston 2018).

What exactly is the free energy principle? It is notoriously difficult to explain, so I will simplify. According to Friston, homeostasis in biology is an expression of an even more basic mechanism in physics, namely, a natural tendency toward "self-organization." Self-organizing systems are intrinsically adaptive: they autonomously maintain their own structural and functional integrity despite environmental disruptions. The principle of self-organization bridges the divide between nonliving systems (like crystals that form spontaneously in liquid) and living ones (like cells), the

[20]Here are two random examples: EEG entropy values are higher in fully conscious than in minimally conscious than in vegetative patients (Gosseries et al. 2011); and baseline fMRI activation increases and decreases predictably when research participants *expect* more or less precision in sensory signals (Hesselmann et al. 2010). Precision is inverse variance, which, like entropy, is a constituent of free energy (as will be explained below).

only difference being their degree of dynamic complexity.[21] Self-organizing systems can become even more complex in the face of external perturbations, as if their interacting components have a unified goal, namely, to maintain their organized state. So self-organizing systems possess properties that Kant (1790) believed impossible: they have *intrinsic* aims and purposes; that is, they are "teleological." Although Darwin never understood it so deeply, "natural selection" is an expression of this tendency. Self-organizing systems arise spontaneously from primal-soup-like conditions, when the short-range interactions between their subsystems become "ergodic"—which just means that they start to occupy limited states. Such systems spontaneously evolve toward a *settling point,* which William Ross Ashby described as an "attractor" in a "basin" of surrounding states. The further evolution of these systems displays some remarkable features. The first is that they acquire a "Markov blanket," which is something like a surrounding membrane that separates the system (and its internal states) from the not-system (its external states). This is the ground-zero origin of *selfhood*. That is, it bestows upon the system a "point of view," which is the ground-zero origin of *subjectivity*. The lawful chains of influence that characterize Markov blankets make it impossible for the internal system to be influenced by the external not-system directly. The system (the "subject") can register its environment (its "objects") only indirectly, *vicariously,* via the effects that external events have on the condition of its blanket. These conditions are called the "sensory states"[22] of the blanket. This sequestered relationship between the interior of the system and the outside world is what generates its point of view. But the sensory states of the blanket have a second function. They influence the inner workings (the "internal states") of the system in such a way that they cause the internal states to generate "active

[21]My point here is not to equate crystals with minds, of course, but rather to note that the *causal mechanism* of self-organization preceded *biological* self-organization, and, indeed, gave rise to it. (In fact, evolution itself is an expression of self-organization; see Friston 2013.) Crystals minimize their free energy in a trivial way because their nonequilibrium steady state has a point attractor. That is, they just arrange themselves into compact patterns and stay there, even when slightly disturbed. Things get far more complex when the attracting set has an itinerant structure with dynamics of the sort that the human brain and mind conform to.

[22]Why are they called "sensory states" rather than "sensations"? This is because these formalisms apply to both living and nonliving systems; as stated above, they bridge the divide between physics and biology and, ultimately, psychology. The same applies to the "active states" described below, which in living systems would normally be called "actions." Note that not all self-organizing systems (even biological ones) are conscious; I have not yet explained how *conscious* selfhood arises.

states" in the blanket, which then influence the external states. The external states, in turn, as we know, influence the sensory states of the blanket. In this way, the consequences of the system's actions are *fed back* to the system's interior. The next remarkable development is that the internal states come to display what can only be described as the ground-zero origin of *learning*. That is, the internal states of the system naturally come to *model* its external states, so that the system generates the active states that are most likely to preserve its structural integrity. This is the crux of self-organization; if the system did otherwise, it would cease to exist as a system. This is the ground-zero origin of *intentionality*. Put differently, the system learns a capacity for statistical inference: it acts upon the world in ways that have proven likely to produce the required sensory consequences. For example, if the system is overheating, it acts in the manner most likely to achieve cooler sensory states. (For the details summarized in this condensed paragraph, and for an empirical demonstration of how self-organizing systems displaying active inference arise naturally from primal-soup-like conditions, see Friston 2013.)

This brings us to the crux of the matter. The difference between the sensory states that are *predicted* by the system's internal states to flow from its active states and the sensory states that *actually* flow from them must be minimized. This difference is called "surprisal." Put simply: the system must avoid surprising (unexpected) states if it is going to survive. This is the mechanism of homeostasis all over again. Leaving aside some of the mathematical complexities, free energy is thus a function of *average surprisal*. The gap between the predicted sensory states of a self-organizing system and its actual states, over a given period of time, is measured as its free energy. Because increasing free energy is an existential threat to the system, the system must minimize the *expected* free energy. And this, it appears, is the fundamental mechanism of drive. Drive is a (quantifiable) measure of the demand made upon the system for *more effective* work.

Here comes the oddest part. When it comes to an information-processing system (like the nervous system and its mental apparatus), this means that *the system must minimize the amount of information it requires from the environment*. If that sounds counterintuitive, remember that it just means the system must have fewer questions of its environment; in other words, it must *reduce its uncertainty*. This, in turn, means that it must *improve its internal model* of the world—its knowledge of how to

meet its needs in the world—and adjust its actions accordingly. And that, we now see, is the reality principle.

What all of this means, at bottom, is that drive pressure (affective arousal)[23] must be minimized, and the way we do this is by generating the best predictions we can as to how we might reduce the pressure by meeting our needs in the world. This is the great task of learning from experience. Let me be as clear as possible, since this is so fundamental: drive is a measure of the expected free energy in the mind, energy that must be deployed in effective work, work that requires learning. Learning, in turn, entails improving the system's predictive model of its self-in-the-world, and this is the ground-zero origin of *representation*.[24]

It only remains to say, before we turn to the classification of the drives and their relationship to the body, that the brain mechanisms by which all of this is achieved are quite well understood. The "body-monitoring" structures that register and regulate homeostatic errors are located in the core brainstem and diencephalic nuclei that surround the central canal and lower cerebral ventricles. Perhaps not surprising is the fact that these nuclei are intimately connected with the brain mechanisms for arousal, the core ones of which are the nuclei of the reticular activating system.[25] What would have surprised Freud, however, is that these mechanisms, the very fount of our drives, *are also the fount of consciousness*. Freud wrote:

> What consciousness yields consists essentially of perceptions of excitations coming from the external world and of feelings of pleasure and unpleasure which can only arise from within the mental apparatus; it is therefore possible to assign to the system *Pcpt.-Cs*. a position in space. It must lie on the borderline between inside and outside; it must be turned towards the external world and must envelop the other psychical systems. It will be seen that there is nothing daringly new in these assumptions; *we have merely adopted the views on localization held by cerebral anatomy, which locates the "seat" of consciousness in*

[23]As Donald Pfaff (2005) beautifully explains, increasing arousal is the same thing as increasing information-processing; the more *uncertain* an animal is, the more aroused it becomes. The converse of arousal, in a sense, is habituation. Therefore, the essential task of mental life (of what we psychoanalysts call "ego development") is to *learn* how to minimize arousal. This is Freud's Nirvana principle, but it is the very opposite of deathly.

[24]Here I am using the term "representation" in Freud's sense, which, let us not forget, was grounded in Franz Brentano's conception of *intentionality* (or "aboutness").

[25]These mechanisms—the source-nuclei for dopamine, serotonin, and norepinephrine—are the ones on which most current psychopharmaceuticals exert their "direct influence" on the drives (i.e., on affect).

> *the cerebral cortex*—the outermost, enveloping layer of the central organ. Cerebral anatomy has no need to consider why, speaking anatomically, consciousness should be lodged on the surface of the brain instead of being safely housed somewhere in its inmost interior [1920, p. 24; emphasis added].

The irony is that consciousness-as-feeling is indeed safely housed in the brain's inmost interior. The system *Cs*. is not a cortical system; it arises from the core of the brainstem (see section [5] above). What this quotation makes clear, once more, is that Freud adopted neuroanatomical views that have since turned out to be wrong. We are therefore obliged to correct him in line with subsequent neuroscientific findings. As I have tried to show elsewhere (e.g., Solms 2018a), correcting metapsychology on this score has practical implications for our therapy. To use the terminology introduced in this paper, improving our *predictive model* of how the mind functions enables us to do *more effective work* in it, and that in turn enhances the chances of survival (and reproductive fitness) of our discipline. (For more detail concerning the clinical application of the metapsychological formulations outlined here, see Solms 2018a.)

(C) Classification of the Drives and the Bodily Sources of Drive

There is no definitive taxonomy of the drives in neurobiology today. This is primarily because different scientists use different classificatory criteria. For example, Paul Ekman (Ekman et al. 1987) uses ethological observations of facial expressions and associated behaviors, whereas Jaak Panksepp (1998) uses the experimental effects of deep brain stimulation, electrical and chemical. Therefore, to pick an example more or less at random, these two scientists classify disgust differently. Ekman classifies it as a "basic emotion," whereas Panksepp classifies it as a "sensory affect." (I will explain these terms in a moment, but note that different taxonomies come with different nomenclatures, which can be confusing.) The important thing for our purposes is that both scientists agree that disgust *exists* and that it reliably displays specific physiological and psychological properties. Taxonomies are artificial constructions that we impose upon nature. Perhaps, therefore, how we *classify* the drives doesn't matter quite as much as how we *understand* them.

Here I will follow Panksepp's classification and nomenclature, as they are the most widely used in affective neuroscience today. (Remember, though, that they are not *universally* accepted.)

Panksepp divides the drives into "bodily" and "emotional" types. He further divides the bodily drives into "homeostatic" and "sensory" subtypes. These terms denote the interoceptive versus exteroceptive *sources* of two subtypes of bodily drive. Hunger, thirst, sleepiness, and bowel distention, for example, are interoceptive, whereas pain, fright, affective touch, and disgust are exteroceptive. I do not use Panksepp's term for the interoceptive bodily drives, because, as I have explained in section (A), and as I shall explain further here, *all* drives are homeostatic. The "emotional" drives are distinguished from the "bodily" ones mainly by virtue of the fact that they do not arise from bodily needs so much as from what might be called *object-relational* ones. That is, they arise from biological needs in relation to other mental *agents*. This makes it more difficult for the subject of these drives to learn how to satisfy them, mainly for the reason that volitional agents are less *predictable* than inert objects. The indefiniteness of such distinctions coincides with the artificiality of our taxonomies. However, it is this distinction that makes the "emotional" drives more significant for psychoanalysts (and, indeed, for psychiatrists and psychotherapists in general). Because it is more difficult to learn how to keep a caregiver close and attentive than it is to learn how to defecate or to withdraw from pain, for example, our patients are far more likely to be troubled by their emotional than by their bodily drives. In addition, as we shall see, emotional drives are far more likely to *conflict* with each other than bodily drives are. This, too, makes it more difficult for us to learn how to master them. This does not mean that our patients never suffer from unsatisfied bodily drive demands (witness sleep disorders, eating disorders, unexplained pain), but experience shows that these bodily ailments are ultimately explicable in emotional terms. For example, it is not the case that patients suffering from anorexia nervosa do not know how to eat; rather, eating, for them, has become conflated with emotional drives they cannot master.

It is important to notice, therefore, contra Freud (see section [2] above), that not all drives have a bodily source. What is the bodily source of longing for the attentive presence of an attachment figure, for example?[26]

[26]This is not to deny that emotional drives (like attachment) serve survival and reproductive success *over the long term* or that they are intimately *connected with* bodily needs (e.g., in the case of attachment, see Hofer 1996). My point is that once an emotional drive (e.g., need for proximity of a caregiver) evolves, it exists in its own right; the need for care is not the same thing as the need for nourishment. This has important implications for Freud's theory of "anaclisis."

For the reasons just stated, I will say relatively little else about the bodily drives here, and will focus instead on the emotional ones. Before proceeding, however, some general principles must be outlined that apply to *all* drives. The first is that they are always associated with affects. (That is why Panksepp speaks of homeostatic and sensory *affects;* recall the distinction made in section [A] above: affects are the subjective aspect of drives.) This is important because many people fail to recognise that bodily states like sleepiness and the urge to defecate are both drives *and* affects, yet that is what they are. I stated in section (A) that the defining feature of every drive is that it is affectively *valenced;* that is, it has a pleasurable and an unpleasurable dimension. Hence, fatigue feels bad and rest feels good; a distended bowel feels bad and voiding it feels good. The distinction between "negative" and "positive" drives is thus a misnomer. Anger is a negative drive, but discharging it (e.g., successfully annihilating a rival) feels good. Fear is a negative drive, but escaping the danger and finding safety feels good. The second principle to notice is that each drive feels good and bad in its own *categorical* way. There are a great variety of different pleasures and unpleasures. Thus, the unpleasure of hunger feels utterly different, qualitatively, from the unpleasure of a full bladder. There is a good reason for this: as explained in section (A), feelings are calls to action that tell us how we are doing in relation to a *specific* need. Accordingly, eating will do nothing to relieve your bladder, and urinating will not satisfy your hunger. (Recall a point I made in section [2] about the "confluence" of drives: each drive *must be satisfied in its own right;* it makes no biological sense to say that 3/10 of thirst plus 7/10 of pain equals 10/20 of total drive pressure. This is why affects are—indeed, must be—distinguished qualitatively, as all categorical variables must be. This is the ground-zero origin of *qualia*. I will return to this important point below.) The third principle is that we are equipped with innate predictions as to what must be done to satisfy each drive. As I have noted, however, these reflexes and instincts must be supplemented through *learning* from experience. This is the main task of mental development: learning how to satisfy our drives, and especially the emotional ones, which are so difficult to master.

Two further general principles in relation to this revised conception of drive shed new light on the questions that bedeviled Freud. I will mention them only briefly now and elaborate later. The first is that there are both "appetitive" and "consummatory" pleasures (see the distinction

between "wanting" and "liking" [Berridge 2003]). Had Freud known this, he would not have been so puzzled by foreplay and, more so, by the fact that sexual excitement is pleasurable. Sexual desire is appetitive; orgasm is consummatory. The second is that drives are aroused by both "needs" and "opportunities." Sexual desire, for example, can have both internal and external triggers. To say this is to say the obvious, but had Freud been familiar with the concept of "incentive salience" he might have thought differently about issues like the bodily sources of drive. As I will explain shortly, however, the sources of our emotional drives are the "inmost interior" *of the brain itself.*

When it comes to the emotional drives, it turns out that *seven* of them can be reliably elicited by electrical or (specific) chemical stimulation at exactly the same brain sites, in all mammals, from mice to men. (Many of them can be evoked in birds, too, and some in all vertebrates.) Mammals and birds shared a common ancestor about 200 million years ago, and vertebrates evolved about 525 million years ago; that is how primitive these drives are.

Panksepp put his terms for the drives in full capitals, to distinguish them from colloquial usage—that is, to indicate that he was talking about whole biological systems, not only the feelings. Here is a brief summary of the seven emotional drives (also known as the "basic emotions").[27]

LUST. One is not constantly sexually aroused. Erotic feelings enter consciousness only when sex is *prioritized* over other things, which happens in the context of fluctuating needs and opportunities (as in the case of *all* drives; for the mechanism of prioritization, see section [D] below). When sexual desire is prioritized, one feels it; then erotic feelings *guide one's voluntary actions*. One pays attention to different things when one is sexually aroused compared to when one is fearful, say, and one behaves differently too. In this way, exteroceptive consciousness and voluntary behavior are determined by your drive state; you experience the world differently—you literally bring different experiences upon your head—depending on what you are feeling. That is why, when the need for safety is prioritized, another drive takes over and sexual needs recede from consciousness.

[27]See Panksepp (1998) for the empirical findings on which the following summary is based, and for the bibliographic references. The remainder of this section draws heavily on Solms (2021).

I am starting this account of the drives with LUST not only because it is the preeminent drive in biology, and in the history of psychoanalysis, but also because it is unclear whether LUST should be classified as a "bodily" or an "emotional" drive. Some people even doubt that sexuality is a *need*. This is an excellent example of the difference between (nonconscious) needs and the *feelings* they give rise to. As discussed in section (A), what motivates subjective beings is the pursuit of pleasure, not reproductive success. The same applies to sweet tastes, for example; we eat candy because we *like* it, not because it contains high energy supplies (which is the evolutionary biological reason why sweet things taste good).[28] That is, *we are driven by feelings*. Living creatures need to reproduce, at least on average; that is why the sexual *drive* became subjectively pleasurable, through natural selection—even though most people know nothing about the underlying biological mechanism, and perhaps do not even endorse it.

I say "on average" because not all sexual activity results in reproduction, just enough of it to keep the species going. This exemplifies a central principle concerning reflexes and instincts: *the limited utility of inborn behaviors to meet our emotional needs*. This is the main reason why emotional drives (as opposed to bodily ones) are so difficult to master. In sex, the inborn reflexes and instincts boil down to little more than genital engorgement and lubrication, lordosis (arching the back, which makes the vagina available for penetration), mounting, intromission, thrusting, and ejaculation. Together with these reflexes, stroking the clitoris or penis (which are anatomically equivalent organs) at a certain rhythm produces pleasurable sensations that predict the release of sexual tension, ideally through orgasm, to satiation.[29] These involuntary contrivances do not equip us for the difficult task of persuading other people—especially the particular ones we are attracted to—to comply with our desire to have sex with them. As already stated, the main reason why "emotional" needs are more difficult to meet than "bodily" ones is that they typically involve other agents *who have needs of their own;* they are not mere substances like food and water. To satisfy the sexual drive, therefore, we must supplement our innate knowledge with additional skills, acquired through

[28]But the sensory-affective pleasure in sweet tastes has nothing to do with reproduction, and it is not driven by LUST. As I have already said, there are many different pleasure systems in the brain.

[29]Note that once satiation is achieved, sex is no longer pleasurable.

learning. This fact alone explains the wide variety of sexual activities that we indulge in, alongside the "average" form.

Notice that learning does not *extinguish* reflexes and instincts; it nuances, elaborates, supplements, and then overrules them—but they are still there. The usual neurocognitive mechanism for updating long-term memories ("reconsolidation"; see section [D] below) doesn't apply to reflexes and instincts. That is because reflexes and instincts are not *memories;* they are fundamental dispositions "hard-wired" into each species through natural selection (see Ellis and Solms 2018).

Our range of sexual behaviors is increased further by the fact that the brain circuits for both female-typical and male-typical LUST exist in every mammal. This is important confirmation of Freud's notion of innate "bisexuality." The gender that comes to dominate in each individual is determined by various factors, centrally including genetic and intrauterine ones (LeVay 1993). Psychoanalysis historically underestimated these latter factors. I will not go into the anatomical and chemical details here, except to point out that both male and female circuits arise in the hypothalamus and terminate in the periaqueductal gray. In other words, like all the basic emotional drive circuits, they are almost entirely subcortical.

The remarks I have made already about the "sources" of the sexual drive may be complemented by a comment on its "objects." Given the difficulties associated with persuading other people (especially preferred people) to have sex with us, why do we not simply satisfy the sexual drive through masturbation?[30] The answer has everything to do with the fact that we have multiple emotional needs. Sexual needs must be met *in conjunction* with all the other needs, such as the "attachment" needs to be discussed below.

SEEKING. All the bodily needs (and, importantly, this includes sexual ones)—which are registered by "need detectors" located mainly in the medial hypothalamus—activate this second emotional drive.[31] The

[30]Note in this respect that the clitoris and penis are not childish "sources" of the sexual drive, as Freud (1905) assumed; rather, they are its first "objects," as he implicitly realized when he developed the theory of narcissism, via "auto-erotism." The erogenous zones are only "sources" of sexuality in the sense mentioned before, that is, of "incentive salience." These zones when stimulated represent "opportunities" rather than "needs." The true sources, in Freud's sense, of LUST are the hormones and peptides testosterone, estrogen, progesterone, vasopressin, oxytocin, LH-RH, and CCK, and these can be aroused both by opportunities and by needs. The same principle applies to all the emotional drives, the homeostatic "control center" of which, in every case, is to be found *in the brain itself*.

[31]These need detectors are the "sensory terminal organs" that Freud (1940) imagined were located in the interior of the body, which he assumed were "offshoots of the cortical layer."

SEEKING drive behaves in almost exactly the same way as Freud's "libidinal" drive, but he did not know that LUST merely *activates* this drive; LUST and SEEKING are not the same thing. SEEKING generates energetic "foraging" behavior, accompanied by a conscious feeling state that may be characterized as expectancy, interest, curiosity, enthusiasm, or optimism. Unlike orgasm, therefore, which is accompanied by a "consummatory" affect, foraging is accompanied by an "appetitive" affect *that is pleasurable in its own right*. This form of pleasure motivates the organism to go out into the world in order to remove (or at least minimize) the unpleasurable tensions generated by its other needs.

Think of a dog in an open field: no matter what its current bodily needs may be, foraging propels it to engage positively with the environment, so that it might satisfy them there. Almost everything that living creatures need is "out there"; then through foraging they *learn,* almost accidentally, which things in the world satisfy each of their needs. In this way, through SEEKING, they encode cause-and-effect relationships and thus bestow value (literally, *meaning*) upon the world. Objects do not, of course, have inherent meaning; their meaning is subjectively conditioned. This illustrates very well how stereotyped "objectless" drives lead to individualized learning from experience. In this respect, the critics of Freudian drive theory cited at the outset of this paper drew a false dichotomy between our "biologized" metapsychology and a supposedly pure psychology of "meaning making."

It is important to recognize that the other drives become "confluent" in the SEEKING drive, but not in the manner Freud thought. First, the other drives are not "components" of SEEKING; rather, they are *triggers* of SEEKING. That is, they exist in their own right; and SEEKING is a drive of its own. Second, SEEKING has no *special* relationship with LUST. In fact, had Freud recognized that LUST activates SEEKING just like all the other drives do, he probably would not have described this energy as "libidinal" or "erotic." It is an *all-purpose* pleasure-seeking (and object-seeking) drive. Third, it is incorrect to suggest that one "component" SEEKING drive can be satisfied by another. All biological needs (including emotional ones) activate SEEKING, the satisfaction of which *leads to* the satisfaction of the initiating drives, depending on the *opportunities* a foraging animal encounters. But each of the underlying needs still must be met in own right, or the animal will die. Finally, LUST does not "lean upon" the self-preservative drives in the sense of Freud's anaclitic

mechanism; rather, LUST and the self-preservative drives, combined, lean upon SEEKING, and SEEKING is a drive *in its own right*.

In this last respect, SEEKING is unique among the drives in that it *proactively* engages with uncertainty (so as to make the world more predictable, in advance, before it surprises us; surprises can be dangerous). This is the origin of novelty-seeking, sensation-seeking, and risk-taking behavior. Foraging makes animals (and people) explore interesting things, so that they know what to expect when they encounter them in the future. Once a dog has explored a hedge, for example, and familiarized itself with its contents, it will be less interested in it the next time around. Accordingly, SEEKING is the "default" drive. When we are not in the grip of one of the other "task-oriented" drives, our consciousness tends toward this generalized interest in the world. This is what Melanie Klein termed an "epistemophilic" drive. It explains why you are reading this paper.

Anatomically, the pathways of the SEEKING circuit arise from the ventral tegmental area of the brainstem, from whence they course upward, to the lateral hypothalamus, nucleus accumbens, amygdala, and medial frontal cortex, and downward (like *all* the drives; see section [D]) to the periaqueductal gray. Chemically, its command neuromodulator is dopamine. Unlike the other reticular activating systems, dopamine activity does not decrease with the onset of sleep. This reveals an interesting fact about SEEKING: it can be aroused even during sleep, by endogenous demands made upon the mind for work, leading to dreams (see Solms 2000). The demands in question need not be bodily ones; offline default-mode memory consolidation, too, can result in prediction errors requiring "mental work" for their resolution (and it is no accident that this work entails conscious feeling, which punctuates the unconsciousness of sleep).

The link with dreams makes clear why excessive SEEKING leads to psychosis, a central mechanism of which may be the forging of *too many* cause-and-effect relations (i.e., too much "meaning making"; see Kapur 2003).[32] It is easy to see why the SEEKING drive is also deeply implicated in addiction (think of cocaine[33] and amphetamines, which stimulate SEEKING directly); it is literally the brain's "craving" system (see

[32]Finding too many cause-and-effect relations arises from excessive SEEKING partly through an inflated sense of agency (megalomania), i.e., an excessive sense of being able to *cause* things.

[33]It is plausible that Freud's discovery of the psychological properties of the "libidinal" drive was facilitated by his own experience of the effects of cocaine.

Volkow, Wise, and Baler 2017). It also plays a big part in gambling and other forms of excessive optimism. Later we shall see that SEEKING plays a pivotal role in mood disorders, as well. Depression is the very opposite of SEEKING (and of mania).

RAGE. While animals engage positively with the world through SEEKING, in the optimistic belief that their drives will be satisfied there, things do not always go well for them. Just as evolutionary prehistory equipped them with reflexes and instincts that reliably predict ways to meet their bodily needs (including foraging), so too they are born with emotional behaviors that predict ways to get them out of trouble. In challenging situations of universal biological significance, these instinctual behaviors are prioritized. We animals are thereby spared the biological costs of having to reinvent the wheels that enabled our ancestors to survive and reproduce. Instincts are a precious inheritance. They transmit innate survival skills—implicit "knowledge" of how to respond in evolutionarily predicted situations—along with the feelings that guide voluntary actions in *unpredicted* ones.

When the RAGE drive is triggered—as it is by anything that gets between an animal and the things that could satisfy its other drives—its consciousness is colored by feelings ranging from irritated frustration to blind fury. (Think of a vending machine that fails to deliver an item you just paid for.) But RAGE is not the only form of aggression. Affective neuroscientists speak also of "cold aggression" (mediated by the SEEKING drive; think of a lion chasing a springbok) and "dominance behavior" (mediated by the PLAY drive, to be described below, which entails a *symbolic* form of aggression: more bark than bite). The form that is mediated by RAGE is called "hot aggression." The reflexes and instincts involved in it include piloerection, protrusion of nails, hissing, growling, and baring of teeth, followed by "affective attack": lunging at the target of wrath and biting, kicking, or hitting it until it relents.[34]

Why do you *feel* the drive pressures that accompany such behavior? The answer is: the feelings tell you *how you are doing,* whether things are going well or badly *for you,* as you try to rid yourself of the obstacle—one that is often simultaneously trying to get rid of you. You sense the sweetness of victory or the bitterness of defeat. This guides what you do next,

[34]This is not an internal death drive directed outward; it is straightforward self-preservative aggression. However, aggression can be directed to the wrong object, and indeed toward the self.

including the possibility that pain (a sensory affect that is suppressed during affective attack) might become prioritized, thereby replacing RAGE and putting an end to the fight—and perhaps leading to flight.

How could this all go on automatically, without conscious evaluation? This question applies also to the role of affect in *thinking*—a topic I can fruitfully introduce here. As Freud taught us, thinking is "trial" action; the capacity to try things out in imagination; a capacity which, for obvious biological reasons, saves lives. This capacity is not unique to humans, but it is particularly highly developed in us. So let me use a human example. Picture this situation, derived from my own experience: My headmaster is frustrating me, arousing feelings of irritation. The instinctual response is affective attack. Now I *think* about the potential consequences. Instead of lunging at him, therefore, I *inhibit* the instinctual action and *imagine* my range of alternatives; I *feel* my way through them ("signal" affect). Eventually I settle upon a satisfactory solution: when nobody is looking, I deflate the tires of my headmaster's car. In this way, I reduce my RAGE without suffering dire consequences. This illustrates, once again, why innate behavioral stereotypes *must* be supplemented by learning from experience, including the imaginary form of experience called thinking. Patients who fail to do this are in serious trouble. When faced with real-life frustrations, which frequently include *conflicting* needs (in this case, RAGE vs. FEAR), instinctual solutions are not enough. But again, please notice: supplementing instinctual responses through learning does not *erase* them. I decided not to attack my headmaster, but the inclination to do so remained, and would arise again in similar situations in future.

Emotions like RAGE are not "mere" feelings. Emotions play a fundamental role in survival and reproductive success. Imagine the consequences if we didn't stake claims on the available resources and prevent others from taking our share. If we couldn't become frustrated, irritated, or angry, we wouldn't be inclined to fight for what we need, in which case, sooner or later, we'd be dead. It is easy to overlook the biological function of emotion in the civilized conditions under which we live today. But we have only been living like this (i.e., in permanent settlements with artificial laws regulating social behavior) for about twelve thousand years. Civilization is a very recent feature of mammalian existence; it played no part in the design of our brains.

Conscious thinking requires cortex, *but the feelings that guide it don't*. The circuit mediating RAGE is almost entirely subcortical, and,

like all the other drive circuits, it terminates in the brainstem periaqueductal gray.

FEAR. The fight/flight dichotomy shows that affective attack is not always the best way to deal with an adversary. The contextual factors separating them are registered in the amygdala, which mediates both RAGE and FEAR drives.

Most mammals "know" from day one that some things are inherently dangerous. Newborn rodents, for example, freeze when exposed to a single cat hair—although they have never experienced cats and know nothing of their attitude to mice. It is easy to see why this is so; if each mouse had to learn from experience how to respond to cats, that would be the end of mice. (Which is probably how "castration anxiety" evolved.) This is one of a limited number of examples of objects that *do* possess intrinsic meaning for us. Again we see the enormous biological value of the basic emotions.

We humans fear some stereotyped things—perhaps all the objects of the common phobias—and we avoid them by the same instincts and reflexes as other mammals do: freezing and fleeing. These behaviors are facilitated by rapid breathing, increased heart rate, and redirection of blood from the gut to the skeletal musculature. (Hence the loss of bowel control associated with extreme fear.) As with other emotions, the conscious feeling of fear tells one whether one is heading toward or away from safety, and thereby guides voluntary behavior.

An interesting example is provided by the patient SM, who suffered from Urbach-Wiethe disease, a rare genetic condition that results in bilateral calcification of the amygdala. She felt no fear. (See Tranel et al. [2006], who researched SM's *subjective* emotionality.) I have studied a large number of patients like this, as there is an unusually high incidence of Urbach-Wiethe disease in South Africa. I found their dreams to be particularly interesting; they are short, simple, and manifestly wishful, like children's dreams (see Blake et al. 2019). One of the patients we studied, whose husband was unemployed, dreamt "My husband found a job; I was very happy." Another, the mother of a disabled child, dreamt "My daughter could walk; I was very happy." Another, who was too poor to travel, dreamt "I was on a bus, going on holiday; I was really happy." Yet another, whose father had died, dreamt "I saw my father again; I was so happy." These dreams are typical of Urbach-Wiethe sufferers. In light of the close similarity between them and the dreams of children, the role

of the FEAR drive in the development of the "censorship" and the "superego" are promising topics for future neuropsychoanalytic research.[35]

As I said, we humans have a few stereotyped fears. Imagine what would happen if each of us had to learn from experience what happens when we jump off a cliff. That is why we are descended from ancestors who felt disinclined to try. Those that did feel so inclined are not our ancestors because they left no offspring. We have every reason to be grateful for this inheritance.

But then one must learn *what else* to fear. One learns from experience—including thinking—that things other than spider bites and falling from dizzying heights can cause one harm. (This is a further illustration of how innate drives give rise to "meaning making.") Electrical sockets and the shocks they produce, for example, could not have been predicted by evolution, but they are just as dangerous as spiders. Also, one must learn what else *to do* when one is fearful, to supplement the instinctual responses. It is not adaptive to freeze or flee from everything that scares you (think of anxiety disorders), just as it is not adaptive to attack everyone that frustrates you. It should be clear by now what role conscious feelings play in this learning process; they teach us what is working and what isn't in relation to the FEAR drive, before it's too late, and thereby help us stay alive.

Fear conditioning reveals important facts about what is conscious and what is not. One of its special features is "single-exposure learning"; a child need only stick its finger into an electric socket *once* to prevent it from ever doing it again. It is easy to see why; the child was lucky to survive the first time, so why repeat the experience? However, as with all the other biological mechanisms that underwrite emotions, the child doesn't know this in any representational sense; conditioning just happens automatically. This is because FEAR conditioning does not require the involvement of cortex. It can even occur in earliest childhood, during the first three years, before the hippocampus (the cortical structure responsible for "declarative" memory) has matured. This is the biological basis of "infantile amnesia." For this reason alone, just like the famous patient HM (who lacked a hippocampus), many neurologically healthy people fear things without consciously knowing why.

[35]What I have in mind here is the interesting possibility that the superego is constituted not only through projected RAGE but also by FEAR, directly, which inevitably *conflicts* with RAGE.

Cognitive scientists attribute fear conditioning to "unconscious" learning, but that is only because they neglect affect. It is true that many people are unconscious of the reasons why they fear things, but they are only too conscious of the associated feelings. Feeling is all that is required to guide voluntary behavior. This is easily demonstrated using a tachistoscope: if words like "murderer" and "rapist" are *subliminally* associated with face A, and "caring" and "kind" with face B, research subjects will *feel* a preference for face B when they are subsequently required to choose between them, even though they cannot say why. "Gut feeling" is what guides this choice, but feelings easily go unrecognized; so they are described cognitively by words like "guessing."

This explains much of the perplexity that surrounds "unconscious emotions" both in cognitive science and in psychoanalysis. It is not the *affects* that are unconscious so much as the *cognitive* things they are about. As we saw above in relation to thinking, it can certainly help to know what one's feelings are about, but that insight is not essential. In fact, sometimes it is better *not* to think before one acts, not least because thinking takes time, and delay can be biologically dangerous (see section [D] below).

The same applies to fear conditioning. Once we have learned to fear something—especially if we do not consciously know why—the association is well-nigh irreversible. As LeDoux (1996) memorably put it, FEAR memories are "indelible." This reveals important facts about unconscious memory in general, a topic to be discussed in the next section. For now I will mention only that "nondeclarative" (unconscious) memories are *hard to forget,* for the same reason that they're unconscious: they entail *less uncertainty* (i.e., are *more generalizable*) and are therefore less subject to contextual revision. This is how acquired behaviors become automatized. Insofar as the purpose of cognition is to learn how to meet our needs in the world, automatization is the ideal of learning.

PANIC/GRIEF. Separation-distress is different from FEAR. This drive emerges only after we become instinctually attached to a caregiver. Unlike fear conditioning, but for equally good biological reasons, this takes time (up to six months): one instance of nurturant care is not enough to determine whether someone can be relied upon forever.

Let me make explicit something that has perhaps not been made sufficiently clear in this section. *All* these emotional drives are homeostatic. Thus, for example, the "expected state" of RAGE is "nothing is getting

between me and the satisfaction of my needs," that of FEAR is "nothing is threatening my life and limb," and that of PANIC/GRIEF is "my caregiver is available and attentive to me." (These nonconscious expectations are not really formulated in words, of course.) Deviations from such expected states are "error signals" (i.e., affects, heading in deteriorating or improving directions; see Figure 1).

When mammals become separated from their attachment figures, an instinctual sequence unfolds, starting with "protest" behavior and followed by "despair." The protest phase is characterized by feelings of panic, together with distress vocalizations and search behavior. The panic is frequently combined with anger—"where *is* she?!"—which evokes another conflict, this time between PANIC/GRIEF and RAGE. The one drive makes you want to keep your caregiver close to you, always and forever, whereas the other drive makes you want to destroy her. Guilt (a *secondary* emotion) is the typical outcome. This is a good example of how secondary emotions (like guilt, shame, jealousy, and envy) arise. Unlike the basic emotions, they are learned constructs—hybrids of emotion and cognition (see Barrett 2017). Note, then, that the homeostatic emotions—the felt drives—under discussion here do not by any means exhaust the full range of human emotionality.

The despair phase of the separation-distress response is characterized by feelings of hopelessness, of literally "giving up." The standard biological explanation is that if the separated pup's crying and searching do not lead quickly to reunion, then the potential costs of alerting predators to its vulnerable state begin to outweigh the benefits. Also, if the pup wanders too far from home base, its chances of being found when the mother returns are reduced. Thus, on statistical balance, giving up (despite how painful it is) becomes the inherited survival strategy. This is another good example of how we know nothing of the biological mechanisms that gave rise to our instincts; all we know about are the feelings.

Here is a classical description of the separation-distress cascade in human children (Bowlby 1960):

> [Protest] may begin immediately or may be delayed; it lasts from a few hours to a week or more. During it the young child appears acutely distressed at having lost his mother and seeks to recapture her by the full exercise of his limited resources. He will often cry loudly, shake his cot, throw himself about, and look eagerly towards any sight or sound which might prove to be his missing mother. All his behaviour suggests strong expectation that she will return. Meanwhile he

> is apt to reject all alternative figures who offer to do things for him, though some children will cling desperately to a nurse.
>
> [In despair, which] succeeds protest, the child's preoccupation with his missing mother is still evident, though his behaviour suggests increasing hopelessness. The active physical movements diminish or come to an end, and he may cry monotonously or intermittently. He is withdrawn and inactive, makes no demands on people in the environment, and appears to be in a state of deep mourning [p. 90].

The latter state is of course akin to depression, which is often accompanied by guilt. As Freud (1917) taught us, depression ("melancholia") is a pathological form of grief ("mourning"). Accordingly, Panksepp and others applied his elucidation of the brain mechanisms of the PANIC/GRIEF drive to developing new treatments for mood disorders (see Yovell et al. 2016; Coenen et al. 2019). Chemically, the transition from "protest" to "despair" is mediated by mu and kappa opioids respectively, the latter of which shut down dopamine (for the immediate effects of which, see the startling case report of Blomstedt et al. [2008], of a patient who fell into a suicidal depression within seconds of a source nucleus for dopamine being electrically suppressed). That is why depression is characterized by the mirror opposite of the feelings that characterize SEEKING (see Solms and Panksepp 2010). GRIEF, at the level of drive theory, is shut-down SEEKING. By contrast, full-blast SEEKING is mania, before it tips into psychosis. The anatomical trajectory of the PANIC component of the PANIC/GRIEF circuit descends from the anterior cingulate gyrus to the periaqueductal gray. Mayberg et al. (2005) and Coenen et al. (2019) show that deep brain stimulation of the PANIC and SEEKING (i.e., inverse GRIEF) circuits can dramatically alleviate symptoms of depression, within seconds, even in the most chronic and treatment-refractory cases. (I am not recommending this treatment but just saying that it provides proof of concept.)

It is interesting that the mu opioid–mediated PANIC system evolved from the brain's older analgesic system; the mental anguish of loss is an extension of the bodily mechanisms for sensory pain. (That is presumably why mental pain is so frequently somatized.) This is a good example of the seamless transition that exists in nature between lifesaving sensory affects and emotional ones. There is nothing ephemeral about emotions. The painful feelings associated with separation and loss—coupled with learning from experience—play a causal role in ensuring the survival of

us mammals (and birds), who *need* caregivers. This continues to apply beyond childhood, of course: the brain circuits just described mediate attachment bonds throughout life, as they also mediate, sadly, various forms of addiction (think of the opiate crisis). Most drugs of abuse act on either the PANIC or the GRIEF (shut-down SEEKING) drives. They are of course forms of self-medication. As we have known since Kraepelin, opiates are excellent (but dangerous) antidepressants.

It is important to acknowledge that Fairbairn and Bowlby, and some other early object-relational theorists, were right about "attachment" being an innate drive (and instinct) entirely separate from sexuality. The evidence for this is now overwhelming. The only link between the LUST and PANIC/GRIEF drives is the one I have mentioned: namely, that they *both* must be satisfied, which means they must be reconciled with each other, which can be difficult. It is difficult not only because the sexual drive is most securely satisfied in a relationship, but also because one's original attachment figure is usually one's mother. That is why it is difficult to reconcile LUST with what Freud called "affectionate bonds."

CARE. This is the flip side of PANIC/GRIEF attachment: we not only need loving care ourselves, but also need to look after others, especially little ones, and especially our own offspring. The so-called maternal instinct exists in all mammals, but not to the same degree, because it is mediated by chemicals found at higher levels (on average) in females: estrogen, prolactin, progesterone, and oxytocin—all of which rise dramatically during pregnancy and childbirth. (Think of post-partum depression.) Also noteworthy is the overlap between the brain chemistry and circuitry for the CARE, PANIC/GRIEF, and female-typical LUST drives. These facts alone readily explain why depression is so much more common (almost three times) in women than in men.

Approximately 80 percent of human females somehow know from childhood that it is "good" to cradle babies to the left of the body midline, whereas males discover this (instinctually) only after they father children (Pileggi, Malcolm-Smith, and Solms 2015). On the other hand, even completely inexperienced boys usually know what to do when a baby cries. They do not prod it with their fingers or pick it up by the foot to see if that helps; they just know (they innately predict) that a "good" thing to do is to hold it close and rock it while making soothing noises.

And yet, as every parent learns, this is not enough. Successfully raising an infant to maturity requires a lot more than instinct. Therefore, as

with the other emotional drives, one must *learn from experience* (including thinking) what to do in the myriad unpredicted situations that arise. As with the other drives, also, decisions in this respect are guided by *feelings* (of care and concern) that tell us whether things are going well or badly. Another reason why a nurturant drive is not enough is that we do not feel *only* love toward our children, as any parent will attest. The resultant conflicts can be resolved only through hybrid cognitive-emotional processes.

Learning how to reconcile the various emotional drives with each other in context-dependent ways is the bedrock of mental health and maturity. This conclusion is more consistent with the evidence now available than the notion of "stages of libidinal development."[36] Consider sustainable romantic partnerships, for example, which require a judicious integration of LUST with childlike PANIC/GRIEF-type attachment (think of the Madonna-whore syndrome), which in turn is difficult to reconcile with the roving SEEKING drive (think of the thrill of novelty), as well as the inevitable frustrations that provoke RAGE (hence the ubiquity of domestic strife), which in turn conflicts with the concerns of nurturant CARE, and so on. Sustaining long-term relationships is just one example of the many challenges that face every human being. To manage these things—to manage life's problems—we use feelings as our compass. It is feeling that guides all learning from experience, in the various forms I have outlined. But biology provides one further drive to help us on our way.

PLAY. We *need* to play. This activity is not as frivolous as it appears: PLAY is the medium through which territories are claimed and defended, social hierarchies are formed ("pecking order"), and in-group and out-group boundaries are forged and maintained. I believe that this drive, more than any other, underpins the oedipus complex (see Solms in press).

People are often surprised to learn that PLAY is a biological drive, but all juvenile mammals engage in vigorous rough-and-tumble. If deprived of their quota on one day, they will try to make it up the next day, as if by rebound. We all know what rough-and-tumble play is, though the form it takes varies slightly from one mammalian species to another. A

[36]Thus Freud's notion of drive "fusion" and "de-fusion" remains serviceable, at least descriptively, if not mechanistically.

play session starts with an "invitation" posture or gesture; then, if the invitation is accepted, the game is on. The one animal or child exuberantly chases the other, they then stop as they wrestle or tickle each other, taking turns to be on top—accompanied by peals of laughter, or the equivalent vocalization depending on the species (even rats "laugh"; see Panksepp 2005). Then they are back on their feet again, chasing each other in the reverse direction. The associated feeling state is equally universal: it is called fun.

Children just love to play. Empirically, however, the majority of play episodes end in tears. This provides an important clue as to what PLAY is all about, biologically speaking; it is about *finding the limits* of what is socially tolerable, acceptable, and permissible. When play is no longer fun for a playmate, often because you are not being "fair," the playmate won't play anymore. A limit has then been reached. The marking of such boundaries is crucial for the formation and maintenance of stable social groups. And the survival of a group is important for the survival of each member of the group in social species such as ours.

A major criterion in this respect is *dominance*. In any play situation, one of the participants takes the lead role, and the other is submissive. This is fun for both parties, so long as the dominant one does not insist on calling the shots *all* the time. The acceptable ratio of turn-taking seems to be somewhere between 60:40 and 70:30. The "60:40 rule" states that the submissive playmate continues playing so long as sufficient opportunity is given to turn the tables.

This reveals a second function of the PLAY drive: namely, the establishment of social hierarchies. Rough-and-tumble play accordingly gives way (especially from puberty onward) to more organized, frankly competitive games. Of course, play is not limited to games of the rough-and-tumble variety. We humans engage in pretend play, in which the participants try out different social roles (e.g., Mother/Baby, Teacher/Pupil, Doctor/Patient, Cop/Robber, Cowboy/Indian, King of the Castle / Dirty Rascal—note the ever present hierarchies). We do not know what goes on in the imagination of other mammals while playing, but we may confidently hypothesize that they too are "trying out" different social roles, and thereby learning what they can and can't get away with.

This suggests a third biological function for PLAY. It requires us (and conditions us) to take account of the feelings of *others*. If we don't, they will refuse to play with us, and then we will be deprived of the enormous

pleasure it yields. The bully might get to keep all the toys, but he will be deprived of all the fun. This, it seems, is why the PLAY drive evolved (and why so much pleasure attaches to satisfying it): it promotes viable social formations. It is, in a word, a major vehicle for the development of *empathy*.[37]

Play episodes come to an abrupt end when they lose their "as if" quality. If I lock up my little sister and throw away the key, then not only have I broken the 60:40 rule but I also am no longer playing the *game* of Cops and Robbers; instead I am imprisoning my sister. In other words, what is governing our mutual behavior now is FEAR or RAGE rather than PLAY. The same applies to the other games enumerated above. Playing "Doctor/Patient" is a game until it becomes real sex; then it is governed by LUST.

The same applies to the oedipus complex. The incest taboo ensures that it remains "as if"; failure to respect this limit has catastrophic consequences for emotional development. Patients who cannot PLAY (i.e., who have not learned how to satisfy this drive effectively, as occurs in character pathology and some perversions, for example) are in serious trouble. They also find it difficult to use psychoanalytic treatment, which is at bottom a form of PLAY (think of the hierarchy, the mutuality, the boundaries, the "as if" quality, etc).

The fact that PLAY hovers, as it were, between all the other emotional drives—trying them out and learning their limits—is perhaps the reason it has not been possible to identify a command neuromodulator for it. Probably it recruits them all. Anyone who doubts that PLAY is definitively a drive, though, should read Pellis and Pellis's *The Playful Brain* (2009).

We don't always like to recognize that humans, like other mammals, *naturally* claim territories and form social hierarchies with clear rules. (The rules governing primate behavior are remarkably complex.) The structure of families, clans, armies, even nations—almost any social group—is undeniably hierarchical and territorial; and this has been so throughout history. The higher the social status of an individual within the group, the greater the access that individual has to the resources in the territory the group controls. This observation is not a matter of personal preference; it is a matter of fact. If we do not face such facts, we cannot

[37]The development of empathy is therefore by no means an automatic process, as the "mirror neuron" theory might suggest. Empathy is not a reflex; it is a developmental achievement (see Solms 2017a).

begin to deal with them. The fact that emotional drives exist does not mean we have no control over them—that we are obliged to bow before "the Law of the Jungle"; but we ignore them at our peril.

It is easy to see how PLAY, in particular, gives rise to social rules. Rules regulate group behavior, and thereby protect us from the excesses of our individual drives. It is also easy to see how social rules encourage complex forms of communication, and how they contribute to the emergence of symbolic functioning in general. The "as if" quality of play suggests that it might even be the biological precursor of thinking (i.e., of virtual vs. real action). Some scientists also believe that *dreaming* is nocturnal PLAY, i.e., trying out the instinctual emotions in an "as if" world. Interestingly, in REM behavior disorder, where the motor paralysis that normally accompanies dreams is lost due to midbrain damage, patients (and experimental animals) physically *enact* the various instinctual stereotypes—e.g., fleeing, freezing, predatory pouncing, affective attack. It cannot be without meaning that RAGE is far more common in humans with this condition than it is in other animals (e.g., cats). Like chimpanzees, humans are very aggressive creatures.

I cannot emphasize enough the clinical utility of the information summarized in this section, both for psychiatric nosology and for the psychodynamic formulation of individual cases. I strongly encourage readers to consult my paper in *Frontiers in Behavioral Neuroscience,* an open-access online journal (Solms 2018a). That paper, too, however, is just a summary; it does not do justice to the richness of what my colleagues and I have learned in several neuropsychoanalytic study groups that for more than six years now have been meeting around the world, discussing our ongoing analytic work with literally hundreds of cases. Our findings will be reported in due course, in a series of clinical papers (meanwhile, see Smith and Solms 2018).

(D) The Relationship between Drive and Conscious Qualities in General

Having proclaimed that drives are conscious, I will conclude with only a brief consideration of the unconscious—since I have written extensively about this topic elsewhere (Solms 2017b, 2018a, 2021). In short, I believe it is a mistake to conflate drives with the unconscious; the "id" and "the unconscious" are not the same thing (see Solms 2013).

I have said that needs remain unconscious (in some cases, nonconscious) until they become drives. Here everything pivots on the fact that we have various competing needs that must be *prioritized*. A strong clue as to

where the selection is made is provided by the fact that *all* needs (both bodily and emotional ones) deliver their residual error signals to the midbrain periaqueductal gray (PAG). The term "residual" here refers to the fact that we are equipped with automatic mechanisms that can meet these needs *up to a point*. These mechanisms include not only the (nonconscious) autonomic reflexes I have discussed but also the (unconscious) behavioral stereotypes that are *acquired* in the manner I will now describe.

The PAG truly is the "inmost interior" of the brain. It surrounds the central canal that joins the third and fourth ventricles and is an extremely primitive structure. Immediately adjacent to it is the tectum, which contains the superior and inferior colliculi. These, in turn, contain layered representations of all the sensory modalities (except smell), which register "where things stand now" in the outside world, in a highly condensed form known as an attentional "saliency map." Thus, this ancient core of the brainstem constantly receives information about the current *internal* conditions in the PAG (i.e., current needs) and the current *external* conditions in the superior colliculi (i.e., current opportunities). The task of this region, then, which also includes the final path to motor output (the midbrain locomotor region) is to make *decisions* about "what to do next." That is why Merker (2007) calls it the "midbrain decision triangle."

Panksepp calls it the primal SELF because the interface I have just described is, in effect, the brain's *meta-homeostat,* which must choose among the competing error signals arriving from all the brain's individual homeostats. Needs must be prioritized because you cannot do everything at once (e.g., eat and sleep). The selection is made on the basis of salience: the most "salient" need is the one that enjoys—or demands—the greatest current opportunity to reduce free energy (as defined in section [B]; for details see Solms 2021). The prioritized need is then *felt,* or, to put it differently, the "decision" emanating from the midbrain decision triangle is announced in the form of a feeling. In other words, *prioritized needs become drives*. This has two consequences. First, the prioritized feeling is the one that is used to guide the next cycle of *voluntary* behavior; and second, all other current needs are relegated to *automatic* control—that is autonomic (internal) control and automatized (external) behavior.

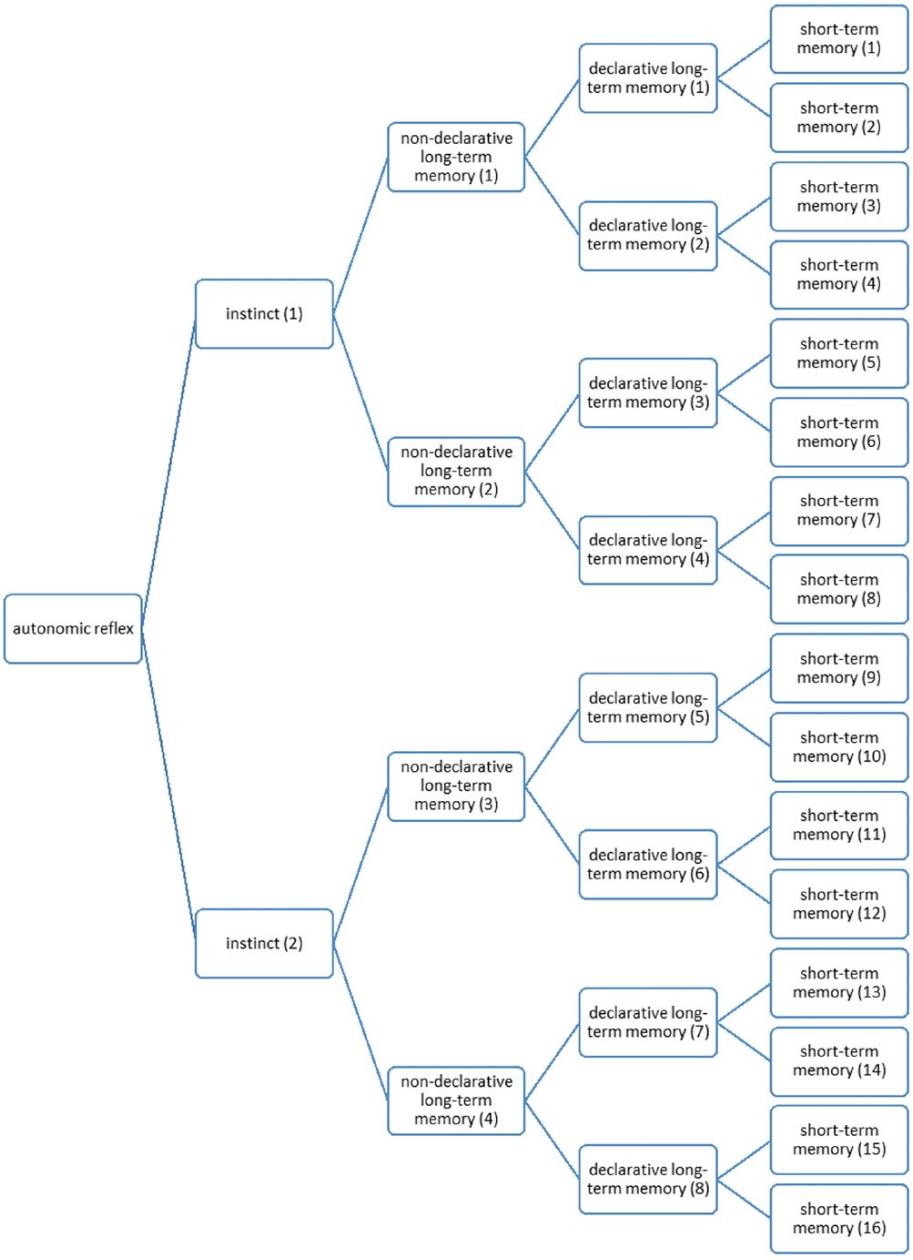
Now the forebrain kicks in. This contains the long-term memory structures that represent a self-organizing system's *predictive model* of the world (see Figure 2). This model is arranged hierarchically, and concentrically, from the autonomic core (which contains innate predictions)

upward and outward, through the subcortical forebrain to the cortex (which contains acquired predictions). This hierarchy becomes *activated* (or, to use Freud's terminology, "cathected"). What is activated, *represented* from long-term memory, is the *expected* sensorimotor context within which the prioritized drive demand will be met. We do not go naively into the world in which we meet our needs; we enter it with expectations (predictions) based not only in reflexes and instincts but also in all that we have learned about how to meet our needs *more effectively* over our individual lifetimes (through learning from experience).

So an expected context is activated in the forebrain by the reticular activating system, which receives massive inputs from the midbrain decision triangle. The forebrain memory traces that represent the expected context are thereby assigned what computational neuroscientists call "precision" weightings and what neurophysiologists call "post-synaptic gain." Precision (or gain) is a measure of *uncertainty,* which, as we saw in section (A) is isomorphic with *arousal*. These precision values—which actually determine the ensuing behavior—are assigned in two different ways. Those pertaining to the automatic behaviors are assigned *fixed* weightings, and those pertaining to the voluntary behaviors (the ones associated with the prioritized drive) are assigned *variable* ones. What this means is that the latter weightings are *modulated* on the basis of the unfolding experience. In other words, they are *palpated by feeling*—which generates cognitive judgments, on the basis that it is "good" when the precision (the uncertainty) attaching to current predictions increases and "bad" when it decreases. In other words, cognitively no less than affectively, *it is good when things turn out as expected and bad when uncertainty prevails*. Increasingly or decreasingly precise predictions, in turn, are determined by the strength of the incoming error signals. In other words, increasingly strong error signals decrease confidence (precision) in the current prediction—which is of course bad from the viewpoint of the system. These error signals are then propagated inward through the predictive hierarchy and they *adjust the predictions that gave rise to them*. This happens iteratively, while progressively better predictions damp down the error signals. Then the whole process begins again, as the residual errors are fed back to the midbrain decision triangle, which then either maintains or shifts the currently prioritized need and associated actions. That, in essence, is how learning works: it entails updating of the predictive model on the basis of strong (i.e., precise) error signals.

Figure 2. A simplified predictive hierarchy, extending from the autonomic core to the sensorimotor periphery

(Learning flows from right to left in this diagram; predictions flow in the opposite direction.)



The constellation of representations that are palpated in the way I have just described are held in short-term or "working" memory. This is an entirely cortical form of representation that is synonymous with cognitive consciousness (i.e., perception and thinking). What is especially

interesting to psychoanalysts is the fact that memory traces activated in this way become *labile* once more; in other words, they literally dissolve and *are memories no longer*.[38] This is because they are in the process of being updated. Neurophysiologists call this updating process "reconsolidation." I have proposed the phrase "predictive work in progress" to describe what reconsolidation does.

It is clear that reconsolidation occurs constantly at the sensorimotor periphery (righthand side of Figure 2); but as a prediction proves increasingly reliable over time, so it slowly becomes more deeply "consolidated" into long-term memory. There are many reasons for this, not least of them being the severe capacity constraints of short-term memory, but the ultimate reason is dictated by the free energy principle itself: the predictive model must be *accurate*. However, accuracy comes with *complexity* costs. Excessive complexity leads to statistical "overfitting" and many other computational and biological problems. The way the system resolves this dilemma is to consolidate predictions that prove to be accurate in a wide variety of contexts (the simpler, more generalizable ones) ever more deeply into the system, ultimately into the "nondeclarative" layers of the predictive model, and to retain only those that must be contextually updated (the more complex, less generalizable ones) in the "declarative" layers. This is because precision modulation has everything to do with changing contexts. The cortex specializes in contexts, and therefore in the *delay* that is the downside (or one of the downsides) of complexity. Thus, nondeclarative predictions are more *certain* than declarative ones, and they can therefore be executed *automatically*. That, I have proposed, is how "the unconscious" (as opposed to the nonconscious) is constituted (Solms 2017b).

Declarative (cortical) long-term predictions are "preconscious," and nondeclarative (subcortical) ones are "unconscious." The unconscious—more certain—ones apply over a wider range of spatial and temporal contexts, which coincides with Freud's description of them as "timeless." These are the predictions that can be executed automatically, in the manner described above, and they behave the most like instincts. It will be noticed that fixed precision values permit what Freud called "freely mobile" cathexis (i.e., no delay), while variable precisions that are "held

[38]See Freud's felicitous phrase "consciousness arises instead of a memory trace" (1920, p. 25).

in mind" (in working memory) during predictive work require what Freud called "bound" cathexis (see Solms 2020). This coincides with his distinction between the primary and secondary processes.

The "repressed" portion of the unconscious consists of predictions that were *prematurely* or *illegitimately* automatized, either because they were automatized before the maturation of the declarative memory systems (infantile amnesia) or because they were the child's "solutions" to problems that proved insoluble. The sexual strivings of children are a good example of an insoluble problem, so their "solution" is almost inevitably subject to repression. Importantly, unlike the other automatized predictions in the "descriptive" unconscious, the repressed ones give rise to *prediction errors* (negative affects; i.e., drive demands). Moreover, unlike declarative predictions, which enjoy updating (reconsolidation) when they give rise to such errors—that is, when they become salient—repressed predictions remain "timeless." They also display the other three "special characteristics of the system *Ucs*." identified by Freud (1915b, p. 187): they show "exemption from mutual contradiction" (which is just another way of saying they are unrealistic and do not actually resolve the problem); they show "replacement of external by psychical reality" (which is another way of saying that the prediction is maintained regardless of the evidence); and they use "primary process" (i.e., freely mobile cathexis). It should be noted that instinctual predictions, too, are indelible, for the reason I have outlined (viz., they are not really memories and so are not subject to reconsolidation). In this sense we can say, with Freud, that instinctual predictions form the "nucleus" of the dynamic unconscious.

Since repressed predictions inevitably produce negative affects (the so-called "return of the repressed"),[39] the only recourse available to the subject in this case is "defense." (Note, therefore, that repression is not synonymous with defense.) Accordingly, patients fall ill when their defenses fail; and *our patients suffer mainly from feelings*. The feelings in question reveal *which* drive demand is not being met, which is why it is important for psychoanalysts to acquaint themselves with the "basic emotions" described above. The repressed prediction, in its turn, is repeated (enacted) in the "transference," for the very reason that it cannot be

[39]The repressed prediction itself does not, and cannot, return ("nondeclarative" memories cannot be remembered, by definition); it is the affect that returns when the defense fails.

remembered. Since I have written elsewhere on these more clinical topics, I will break off at this point (see Solms 2018b).

CONCLUSION

In this paper I have proposed extensive revisions to Freud's theory of the drives. The main revisions are as follows: (1) Drives are *conscious;* they are in fact the fount of all consciousness. (2) Drive energy may be equated with Friston free energy and is therefore *quantifiable* in principle. (3) There are not two drives but *many,* seven of which may be described as "emotional" as opposed to "bodily." (4) All drives are self-preservative or preservative of the species; there is no death drive at work in the mind. That is, all drives are homeostatic and therefore anti-entropic. (5) The great task of mental development is to supplement instinctual predictions about how these multiple drive demands may be satisfied *and reconciled with each other* in unpredicted contexts. This work is done through learning from experience, which is governed by the Law of Affect.

Although these revisions are derived from neuroscientific research findings, which implies that they are open to ongoing testing by neuroscientific methods, I want to end by insisting that the *clinical* situation remains best suited to determining and refining the utility of new psychoanalytic theories. When Freud (1915a) wrote that "working over the psychological material . . . seems rather itself to call for the application to the material of definite assumptions concerning the life of the drives, and it would be a desirable thing if those assumptions could be taken from some other branch of knowledge and carried over to psychology" (p. 124), he surely did not mean that a psychological working-over of the material *adds nothing* to the assumptions carried over from biology. The revision of drive theory that I have proposed here, therefore, some years after the theoretical breakthrough that gave rise to it (Solms 2013), is as much a result of that breakthrough as it is of psychological "working over" of it in the clinical study groups I have mentioned.

Although this article is addressed primarily to a psychoanalytic audience, it contains insights about the lived life of the mind, obtained by the psychoanalytic method, that can be "carried over" in the reverse direction, back to neuroscience (see Solms 2021). As Oliver Sacks (1984) once said: "Neuropsychology, like classical neurology, aims to be entirely objective, and its great power, its advances, come from just this. But a

living creature, and especially a human being, is first and last active—a subject, not an object. It is precisely the subject, the living 'I,' which is being excluded. Neuropsychology is admirable, but it excludes the psyche—it excludes the experiencing, active, living 'I'" (p. 164).

REFERENCES

Barrett, L.F. (2017). *How Emotions Are Made: The Secret Life of the Brain*. New York: Houghton Mifflin Harcourt.

Berridge, K. (2003). Pleasures of the brain. *Brain & Cognition* 52:106–128.

Blake, Y., Terburg, D., Balchin, R., van Honk, J., & Solms, M. (2019). The role of the basolateral amygdala in dreaming *Cortex* 113:169–183.

Blass, R., & Carmeli, Z. (2007). The case against neuropsychoanalysis: On fallacies underlying psychoanalysis' latest scientific trend and its negative impact on psychoanalytic discourse. *International Journal of Psychoanalysis* 88:19–40.

Blomstedt, P., Hariz, M.I., Lees, A., Silberstein, P., Limousin, P., Yelnik, J., & Agid, Y. (2008). Acute severe depression induced by intraoperative stimulation of the substantia nigra: A case report. *Parkinsonism & Related Disorders* 14:253–256.

Bowlby, J. (1960). Separation anxiety. *International Journal of Psychoanalysis* 41:89–113.

Cannon, W.B. (1926). Physiological regulation of normal states: Some tentative postulates concerning biological homeostatics. In *Jubilee Volume to Charles Richet*. Paris: Éditions Medicales, pp. 91–93.

Carhart-Harris, R., & Friston, K. (2010). The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas. *Brain* 133:1265–1283.

Coenen, V.A., Bewernick, B.H., Kayser, S., Kilian, H., Boström, J., Greschus, S., Hurlemann, R., Klein, M.E., Spanier, S., Sajonz, B., Urbach, H., & Schlaepfer, T.E. (2019). Superolateral medial forebrain bundle deep brain stimulation in major depression: A gateway trial. *Neuropsychopharmacology* 44:1224–1232.

Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon.

Damasio, A. (2018). *The Strange Order of Things: Life, Feeling, and the Making of Cultures*. London: Penguin Random House.

Du Bois–Reymond, E., ed. (1842). *Jugendbriefe von Emil Du Bois–Reymond an Eduard Hallmann, zu seinem hundertsten Geburtstag, dem 7. November 1918*. Berlin: Reimer, 1918.

Ekman, P., Friesen, W.V., O'Sullivan, M., Chan, A.Y.C., Diacoyanni-Tarlatzis, I., Heider, K.G., Krause, R., Lecompte, W.A., Pitcairn, T., & Ricci-Bicci, P.E. (1987). Universals and cultural differences in the judgements of facial expressions of emotion. *Journal of Personality & Social Psychology* 53:712–717.

Ellis, G., & Solms, M. (2018). *Beyond Evolutionary Psychology: How and Why Neuropsychological Modules Arise*. Cambridge: Cambridge University Press.

Freud, S. (1894). The neuro-psychoses of defence. *Standard Edition* 3:45–61.

Freud, S. (1895). Project for a scientific psychology. *Standard Edition* 1:295–397.

Freud, S. (1905). Three essays on the theory of sexuality. *Standard Edition* 7:130–243.

Freud, S. (1915a). Instincts and their vicissitudes. *Standard Edition* 14:117–140.

Freud, S. (1915b). The unconscious. *Standard Edition* 14:166–215.

Freud, S. (1916–1917). Introductory lecturers in psycho-analysis. *Standard Edition* 15/16.

Freud, S. (1917). Mourning and melancholia. *Standard Edition* 14:243–258.

Freud, S. (1920). Beyond the pleasure principle. *Standard Edition* 18:7–64.

Freud, S. (1923). The ego and the id. *Standard Edition* 19:12–66.

Freud, S. (1925). An autobiographical study. *Standard Edition* 20:7–70.

Freud, S. (1927). Postscript to a discussion on lay analysis. *Standard Edition* 1920: 251–258.

Freud, S. (1940). An outline of psycho-analysis. *Standard Edition* 23:144–207.

Friston, K. (2009). The Free Energy Principle: A rough guide to the brain? *Trends in Cognitive Sciences* 13:293–301.

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface,* September 6.

Gill, M.M., & Holzman, P.S., eds. (1976). *Psychology versus Metapsychology: Psychoanalytic Essays in Memory of George S. Klein*. New York: International Universities Press.

Gosseries, O., Schnakers, C., Ledoux, D., Vanhaudenhuyse, A., Bruno, M.-A., Demertzi, A., Noirhomme, Q., Lehembre, R., Dams, P., Goldman, S., Peeers, E., Moonen, G., & Laurys, S. (2011). Automated EEG entropy measurements in coma, vegetative state/unresponsive wakefulness syndrome and minimally conscious state. *Functional Neurology* 26:25–30.

Hesselmann, G., Sadaghiani, S., Friston, K., & Kleinschmidt, A. (2010). Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PLoS One* 5(3), March 29.

Hofer, M. (1996). Multiple regulators of ultrasonic vocalization in the infant rat. *Psychoneuroendocrinology* 21:203–217.

Holt, R., ed. (1967). *Motives and Thought: Psychoanalytic Essays in Honor of David Rapaport*. New York: International Universities Press.

Jaynes, E. (1957). Information theory and statistical mechanics. *Physical Review* 106:620–630.

Kant, I. (1790). Kritik der Urteilskraft. *Kants gesammelte Schriften* 5. Berlin: de Gruyter.

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry* 160:13–23.

Klein, G. (1976). *Psychoanalytic Theory: An Exploration of Essentials*. New York: International Universities Press.

LeDoux, J. (1996). *The Emotional Brain*. New York: Simon & Schuster.

LeVay, S. (1993). *The Sexual Brain*. Cambridge: MIT Press.

Masson, J.M. (1985). *The Complete Letters of Sigmund Freud to Wilhelm Fliess, 1887–1904*. Cambridge: Harvard University Press.

Mayberg, H., Lozano, A., Voon, V., McNeely, H., Seminowicz, D., Hamani, C., Schwalb, J., & Kennedy, S. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron* 45:651–660.

Merker, B. (2007). Consciousness without cerebral cortex: A challenge for neuroscience and medicine. *Behavioral & Brain Sciences* 30:63–68.

Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press.

Panksepp, J. (2005). Beyond a joke: From animal laughter to human joy? *Science* 308:62–63.

Panksepp, J., & Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. New York: Norton.

Pellis, S., & Pellis, V. (2009). *The Playful Brain: Venturing to the Limits of Neuroscience*. Oxford: One World.

Pfaff, D. (1999). *Drive: Neurobiological and Molecular Mechanisms of Sexual Motivation*. Cambridge: MIT Press.

Pfaff, D. (2005). *Brain Arousal and Information Theory*. Cambridge: Harvard University Press.

Pileggi, L., Malcolm-Smith, S., & Solms, M. (2015). Investigating the role of social-affective attachment processes in cradling bias: The absence of cradling bias in children with autism spectrum disorders. *Laterality* 20:154–170.

Sacks, O. (1984). *A Leg to Stand On*. New York: Simon & Schuster.

Schafer, R. (1976). *A New Language for Psychoanalysis*. New Haven: Yale University Press.

Smith, R., & Solms, M. (2018). Examination of the hypothesis that repression is premature automatization: A psychoanalytic case report and discussion. *Neuropsychoanalysis* 20:47–61.

Solms, M. (2000). Dreaming and REM sleep are controlled by different brain mechanisms. *Behavioral & Brain Sciences* 23:843–850.

Solms, M. (2013). The conscious id. *Neuropsychoanalysis* 14:5–85.

Solms, M. (2017a). Empathy and the problem of other minds: A neuropsychoanalytic perspective and a clinical vignette. In *Empathy: Epistemic Problems and Cultural-Historical Perspectives of a Cross-Disciplinary Concept,* ed. V. Lux & S. Weigl. London: Palgrave Macmillan, pp. 93–114.

Solms, M. (2017b). What is "the unconscious," and where is it located in the brain? A neuropsychoanalytic perspective. *Annals of the New York Academy of Sciences* 1406:90–97.

Solms, M. (2018a). Extracts from the Revised Standard Edition of Freud's complete psychological works. *International Journal of Psychoanalysis* 99:11–57.

Solms, M. (2018b). The neurobiological underpinnings of psychoanalytic theory and therapy. *Frontiers of Behavioral Neuroscience* 12:294.

Solms, M. (2020). New project for a scientific psychology: General scheme. *Neuropsychoanalysis* 22:5–35.

Solms, M. (2021). *The Hidden Spring: A Journey to the Source of Consciousness*. New York: Norton.

Solms, M. (in press). Revision of the theory of the biological origins of the oedipus complex. *Psychoanalytic Quarterly*.

Solms, M., & Friston, K. (2018). How and why consciousness arises: Some considerations from physics and physiology. *Journal of Consciousness Studies* 25:202–238.

Solms, M., & Panksepp, J. (2010). Why depression feels bad. In *New Horizons in the Neuroscience of Consciousness,* ed. E. Perry, D. Collerton, F. LeBeau, & H. Ashton. Amsterdam: John Benjamins, pp. 169–179.

Sulloway, F. (1979). *Freud, Biologist of the Mind*. New York: Basic Books.

Tranel, D., Gullickson, G., Koch, M., & Adolphs, R. (2006). Altered experience of emotion following bilateral amygdala damage. *Cognitive Neuropsychiatry* 11:219–232.

Volkow, N., Wise, R., & Baler, R. (2017). The dopamine motive system: Implications for drug and food addiction. *Nature Reviews Neuroscience* 18:741–752.

YOVELL, Y., BAR, G., MASHIAH, M., BARUCH, Y., BRISKMAN, I., ASHEROV, J., LOTAN, A., RIGBI, A., & PANKSEPP, J. (2016). Ultra-low-dose buprenorphine as a time-limited treatment for severe suicidal ideation: A randomized controlled trial. *American Journal of Psychiatry* 173:491–498.

University of Cape Town
Neuroscience Institute and Psychology Department
Private Bag X3
Rondebosch, Western Cape 7701
SOUTH AFRICA
Email: mark.solms@uct.ac.za