

Linking Mind and Brain in the Study of Mental Illnesses: A Project for a Scientific Psychopathology

Nancy C. Andreasen

Brain research on mental illnesses has made substantial advances in recent years, supported by conceptual and technological developments in cognitive neuroscience. Brain-based cognitive models of illnesses such as schizophrenia and depression have been tested with a variety of techniques, including the lesion method, tract tracing, neuroimaging, animal modeling, single-cell recording, electrophysiology, neuropsychology, and experimental cognitive psychology. A relatively sophisticated picture is emerging that conceptualizes mental illnesses as disorders of mind arising in the brain. Convergent data using multiple neuroscience techniques indicate that the neural mechanisms of mental illnesses can be understood as dysfunctions in specific neural circuits and that their functions and dysfunctions can be influenced or altered by a variety of cognitive and pharmacological factors.

In 1895, a little-known Viennese neuropsychiatrist named Sigmund Freud wrote a largely unnoticed work entitled *A Project for a Scientific Psychology*, in which he proposed that the cognitive mechanisms of normal and abnormal mental phenomena could be explained through orderly and rigorous study of brain systems. Freud began his career researching pharmacology (the therapeutic effects of cocaine), neurology (aphasia in children), and basic neuroscience (staining techniques for visualizing neurons), but he ultimately abandoned both the project and neuropsychiatry. During the *fin de siècle* 1900s, however, Freud's project is slowly being achieved. This fruition reflects the maturity of the techniques of neuroscience, as well as the convergence of efforts from multiple domains: psychiatry, cognitive psychology and neuropsychology, and clinical and basic neuroscience. Models of illness mechanisms have been developed through the use of clinical observation, experimental paradigms developed in psychology, animal and human lesion studies, anatomic studies of neural circuits, neuroimaging, and behavioral neuropharmacology. The long-term goal is to achieve a "scientific psychopathology": to identify the neural mechanisms of normal cognitive processes and to understand how they are injured in mental illnesses.

This overview provides a summary of some of the fundamental conceptual issues that are being addressed in pursuit of this long-term goal. The work of neuroscientists studying two common mental illnesses—

schizophrenia and depression—illustrates the consensus that is developing among investigators who begin with different strategies originating from different disciplines in the broad field of cognitive neuroscience.

Fundamental Conceptual Issues

The relationship between mind and brain. Mental illnesses have historically been distinguished from other medical illnesses because they affect the higher cognitive processes that are referred to as "mind." The relationship between mind and brain has been extensively discussed in contemporary philosophy and psychology, without any decisive resolution (1). One heuristic solution, therefore, is to adopt the position that the mind is the expression of the activity of the brain and that these two are separable for purposes of analysis and discussion but inseparable in actuality. That is, mental phenomena arise from the brain, but mental experience also affects the brain, as is demonstrated by the many examples of environmental influences on brain plasticity (2). The aberrations of mental illnesses reflect abnormalities in the brain/mind's interaction with its surrounding world; they are diseases of a psyche (or mind) that resides in that region of the soma (or body) that is the brain.

Mind and brain can be studied as if they are separate entities, however, and this is reflected in the multiple and separate disciplines that examine them. Each uses a different language and methodology to study the same quiddity. The challenge in developing a scientific psychopathology in the 1990s is to use the power of multiple disciplines. The study of mind has been the prov-

ince of cognitive psychology, which has divided mind into component domains of investigation (such as memory, language, and attention), created theoretical systems to explain the workings of those domains (constructs such as memory encoding versus retrieval), and designed experimental paradigms to test the hypotheses in human beings and animals (3). The study of brain has been the province of several disciplines. Neuropsychology has used the lesion method to determine localization by observing absence of function after injury, whereas neuroanatomy and neurobiology have mapped neural development and connectivity and studied functionality in animal models (4). The boundaries between all these disciplines have become increasingly less distinct, however, creating the broad discipline of cognitive neuroscience. The term "cognitive" has definitions that range from broad to narrow; its usage here is broad and refers to all activities of mind, including emotion, perception, and regulation of behavior.

Contemporary psychiatry studies mental illnesses as diseases that manifest as mind and arise from brain. It is the discipline within cognitive neuroscience that integrates information from all these related disciplines in order to develop models that explain the cognitive dysfunctions of psychiatric patients based on knowledge of normal brain/mind function.

Using the phenomenotype to find the biotype. There are at present no known biological diagnostic markers for any mental illnesses except dementias such as Alzheimer's disease. The to-be-discovered lesions that define the remainder of mental illnesses are likely to be occurring at complex or small-scale levels that are difficult to visualize and measure, such as the connectivity of neural circuits, neuronal signaling and signal transduction, and abnormalities in genes or gene expression. Despite their lack of a defining objective index such as glucosuria is for diabetes, however, these illnesses are very real. Not only do they produce substantial morbidity and mortality, but advances in psychiatric nosology have produced objective, criterion-based, assessment techniques that produce reliable and precise diagnoses (5). In the absence of a pathological marker, the current definitions of mental illnesses are syndromal and are

The author is at the Mental Health Clinical Research Center, The University of Iowa Hospitals and Clinics and College of Medicine, 200 Hawkins Drive, Iowa City, IA 52242, USA.



based on a convergence of signs, symptoms, outcome, and patterns of familial aggregation (6).

Finding the neural mechanisms of mental illnesses must be an iterative process; syndromal clinical definitions (or the phenotypic) are progressively tested, refined, and redefined through the measurement of neurobiological aspects (or the biotype) (7). This process is not fundamentally different from that used to study other diseases. The diagnosis of diabetes, for example, has evolved from the observation of glucosuria to multiple subdivisions based on age of onset, severity of symptoms and complications, degree of islet cell involvement, and genetic factors. For most mental illnesses, the task is simply made more challenging by the absence of an objective criterion that can provide an initial clue to assist in finding mechanisms, as neuritic plaques have done for Alzheimer's disease.

Defining the boundary between normal and abnormal cognitive processes: continua versus categories. Defining a boundary between normal and abnormal cognitive processes, which demarcates the phenotypic of specific disorders, is often considered to be the first step in identifying the neural substrates of mental illnesses. This task has been difficult. Many of the symptoms of mental illness are on a continuum with normality. The dysphoric mood of depressive illness shares many features with the normal sadness experienced as a consequence of a personal loss such as the death of a loved one or the termination of a marriage. At what point does such normal sadness become a form of psychopathology? Most children are rambunctious and have short attention spans. At what point does this pattern become sufficiently severe to diagnose attention deficit hyperactivity disorder? Thresholds based on duration (such as dysphoria that persists for more than a month after a personal loss) or on severity (such as inattentiveness that interferes significantly with school performance) are usually applied to resolve this problem (8). These are boundaries of convenience that permit reliable definition, not boundaries with any inherent biological meaning.

This approach to defining the boundaries between mental illnesses and normality may seem imprecise. Traditional medical thinking teaches that disease processes are discontinuous from normality. One either has cancer or one does not. Categorical disease models are being challenged, however, by the recent data indicating that individuals may carry a genetic risk factor to develop a disorder that can be measured premorbidly (such as the BRCA1 and ApoE lipoprotein genes) and that may or may not ultimately be expressed as the full form of the disorder,

depending on the occurrence of a variety of cofactors (9). Such observations raise questions about the discontinuity requirement for a definition of a disease process. Current models of the etiology of mental illnesses in fact share many features with cancer. In both groups of disorders, genetic factors may play a role in inducing vulnerability, but additional "hits" appear to be needed to produce a full-blown catastrophic condition that is recognized as an illness. Before the development of a full-blown catastrophic condition, however, the boundary between health and disease may be blurry.

Focus on diseases versus symptoms. As scientists studying mental illnesses work to give shape and focus to these apparently amorphous disease processes, they must also determine whether the problem might be made more soluble by examination of symptoms rather than disease categories. Most disease categories are constellations of symptoms. The nature of the constellation may provide a clue to mechanism (as the combination of dysphoria with diurnal dysregulation and appetitive changes suggests endocrine dysfunction in depression). An alternate strategy in the search for neural mechanisms is to reduce the field of view and examine specific components of disease processes, thereby potentially improving the focus of the investigation.

For example, hallucinations are a common symptom of severe mental illnesses. Understanding their neural mechanisms could potentially tell us something about broader aspects of brain dysfunction in these disorders. Hallucinations are defined as sensory experiences (such as seeing objects or people, or having the sensation that insects are crawling under the skin) that occur without any external stimulation. The most common form of hallucination is auditory. Auditory hallucinations are the subjective experience of hearing a voice that is clearly distinct from the person's own thoughts and has the qualities of an auditory perception (such as volume, pitch, and intonation). The auditory hallucinations experienced during a psychotic illness such as schizophrenia are almost invariably experienced as human voices that speak in sentences or fragments of sentences. Because the brain mechanisms that control auditory perception are well understood and those governing language processing are at least partially understood, identifying the brain disturbances that produce this particular symptom could provide some leverage for understanding the nature of the brain injury in schizophrenia. Shifting the focus of investigation from disease category to specific symptoms offers the possibility of making the search for mechanisms more tractable.

This strategy also introduces other chal-

lenges, however. First, hallucinations occur in a variety of sensory modalities. Auditory hallucinations are the most common, but hallucinations may also be visual, tactile, olfactory, or (very rarely) gustatory. Does the same basic process underly all hallucinatory experiences, with some more specific process leading them to be expressed in a given sensory domain? Or is each a different discrete process? The search for the neural mechanisms of hallucinations will be driven differently, depending on which of these hypothetical alternatives is chosen. A single basic process suggests (but does not require) a subcortical locus, whereas a discrete process suggests specific cortical loci.

Second, hallucinations occur in the same modality in different disease processes. For example, auditory hallucinations are very common in mania as well as schizophrenia. Might one more powerfully seek the common basic mechanism of hallucinations by pooling patients who are traditionally studied separately? On the other hand, hallucinations also occur in different modalities in patterns that are highly characteristic of different disease processes. For example, auditory hallucinations are very common in schizophrenia and may co-occur with tactile hallucinations, whereas visual and olfactory hallucinations are much less common. Olfactory hallucinations are common in temporal lobe epilepsy, whereas visual hallucinations are common in a variety of drug-induced psychotic states (such as delirium tremens or LSD or PCP psychosis) but not in all (for example, amphetamine psychosis). The variability in patterning across disease states argues against the likelihood that hallucinations share the same mechanism in different mental illnesses, and it may argue against pooling patients for study who experience hallucinations in the same modality independently of disease process.

Finally, hallucinations have different and characteristic courses in different disease processes. The auditory hallucinations of schizophrenia tend to be relatively chronic and lifelong, whereas those of mania are brief and only occur during the euphoric mood state that defines a manic episode. This observation suggests that neuroanatomic mechanisms exert a more powerful effect in schizophrenia, whereas the mechanism driving hallucinations in mania is more rapidly reversible and plastic, and therefore neurochemical. Both types can, however, be powerfully affected by the same antipsychotic drugs and may therefore share some common neurochemical mechanism.

Characteristics of heuristic cognitive models of mental illnesses. Creating a scientific psychopathology requires the development of heuristic and testable models. Cognitive

models of mental illnesses may develop from a variety of different vantage points, such as neurobiology, neuropsychology, cognitive psychology, or psychiatry. Whatever the point of origin, models converge on a final common pathway that leads to a shared set of characteristics. The models may begin at the level of systems, but they must also be adaptable to smaller biological scales (such as cells, membranes, and molecules), at which the primary lesions of mental illness probably occur.

Cognitive models of mental illnesses that can be heuristically applied on multiple levels share three characteristics. (i) They provide a general theory of the disease that is consistent with our current level of clinical knowledge, based on observation of signs and symptoms, response to treatment, course of illness, and familial/genetic data. (ii) They provide a theory that can be tested experimentally in human beings, preferably with several types of techniques so that comparative confirmatory or disconfirmatory data can be obtained. (iii) They provide a theory that can also be modeled and tested in animals, because animal models offer the most flexible and powerful techniques for rapidly screening new treatments and for identifying molecular mechanisms of illness.

Linking Mind and Brain: The Examples of Schizophrenia and Depression

Advances that have been made in the study of schizophrenia and depression illustrate the power of developing cognitive models that derive from different perspectives and apply techniques from multiple domains.

Finding the common thread in schizophrenia. The name “schizophrenia” (“fragmented mind”) was coined by Eugen Bleuler, who wished to emphasize that it was a cognitive disorder in which the “fabric of thought and emotion” was torn or fragmented, and normal connections or associations were no longer present (10). Schizophrenia poses special challenges to the development of cognitive models because of the breadth and diversity of its symptoms. The symptoms include nearly all domains of function: perception (hallucinations), inferential thinking (delusions), fluency of thought and speech (alogia), clarity and organization of thought and speech (“formal thought disorder”), motor activity (catatonia), emotional expression (affective blunting), ability to initiate and complete goal-directed behavior (avolition), and ability to seek out and experience emotional gratification (anhedonia). Not all these symptoms are present in any given patient, however, and none is pathognomonic of the illness. An initial survey of the

diversity of symptoms might suggest that multiple brain regions are involved, in a spotty pattern much as once occurred in neurosyphilis. In the absence of visible lesions and known pathogens, however, investigators have turned to the exploration of models that could explain the diversity of symptoms by a single cognitive mechanism. Exemplifying this strategy are four different models that illustrate the melding of cognitive neuroscience and psychiatry, beginning at four different points of departure. The convergent conclusions of these different models are striking.

Cognitive psychology. Approaching schizophrenia from the background of cognitive psychology, Frith has divided the symptoms of schizophrenia into three broad groups or dimensions: disorders of willed action (which lead to symptoms such as alogia and avolition), disorders of self-monitoring (which lead to symptoms such as auditory hallucinations and delusions of alien control), and disorders in monitoring the intentions of others (“mentalizing”) (which lead to symptoms such as “formal thought disorder” and delusions of persecution). Frith believes that all these are special cases of a more general underlying mechanism: a disorder of consciousness or self-awareness that impairs the ability to think with “metarepresentations” (higher order abstract concepts that are representations of mental states) (11). Frith and his collaborators are currently testing this conceptual framework using positron emission tomography (PET).

Their earliest efforts focused on willed action, tested by giving people tasks for which the correct response is not evident from context, such as verbal fluency or choosing a finger movement. Normal people activate a frontal circuit during such tasks, whereas patients with schizophrenia show relative decreases in frontal regions and increases in temporal regions in comparison with normals (12). If the pace of the verbal fluency task is slowed, however, frontal function is similar to that of normals and only the temporal abnormality remains (13). Molecular manipulation with dopamine agonists such as apomorphine also enhances the temporal increases (14). Examination of the correlations between blood flow in frontal and temporal regions suggests that the normal relationship between them has broken down and that there is abnormal functional connectivity (15). Figure 1 illustrates the pattern of temporal lobe hyperactivity observed in these studies.

More recently, this group has completed a systematic study of hallucinations. They hypothesize that hallucinations are due to an erroneous attribution of the person’s own inner speech to another person, reflecting a defect in self-monitoring. Starting

first with normals, they developed a task that could potentially mimic this mechanism of hallucinations; people were asked to perform a sentence completion task and imagine that the response was spoken in another person’s voice; they found that this task led to activation of speech production and perception regions, such as Broca’s area, the supplementary motor area, and the left superior and middle temporal regions (16). Applying the same task to people with schizophrenia and comparing hallucinators to nonhallucinators, they found the hallucinators to have decreased flow in the areas used to monitor speech, such as the left middle temporal gyrus and supplementary motor area (17). In other studies, they have examined flow in patients while the patients were experiencing auditory hallucinations, and they found activations primarily in subcortical regions (the thalamus and striatum), limbic and paralimbic regions (the anterior cingulate and parahippocampal gyrus), and cerebellum. They speculate that activity in subcortical regions may generate or moderate hallucinations, whereas the content (auditory or tactile, for example) may be determined by the specific neocortical regions that are engaged (18).

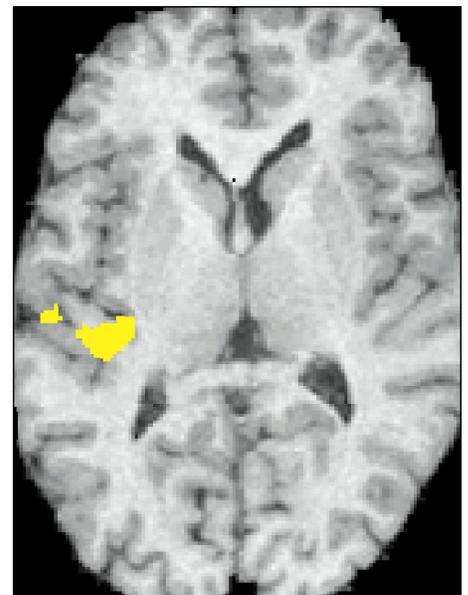


Fig. 1. The area in the left superior temporal gyrus where there is overactivity during the performance of a word generation task is shown in yellow. The data came from 12 drug-free schizophrenic patients (9 of whom were drug-naïve) scanned during an acute phase of their illness (26). The horizontal slice shown is 8 mm above the anterior commissure–posterior commissure line. Significant differences derived from a PET study are shown superimposed on a standard MRI slice in the space of the atlas of Talairach and Tournoux. The location of peak activity is at Talairach coordinates 38, –28, 8.



Recently they have also developed an experimental model for their third symptom dimension: a defect in “mentalizing,” or monitoring the intentions of others. People were given a story comprehension task that required the attribution of mental states to others; during this task, flow was increased in the left medial frontal gyrus and the posterior cingulate (19); at present, no reports have been published concerning performance of this task by schizophrenic patients.

Neurobiology. Approaching schizophrenia from a background that blends lesion studies and single-cell recordings in nonhuman primates for the study of cognition, Goldman-Rakic has proposed a model suggesting that the fundamental impairment in schizophrenia is an inability to guide behavior by representations, often referred to as a defect in working memory (20). Working memory, or the ability to hold a representation “online” and perform cognitive operations using it, permits individuals to respond in a flexible manner, to formulate and modify plans, and to base behavior on internally held ideas and thoughts rather than being driven by external stimuli (21). A defect in this ability can explain a variety of symptoms of schizophrenia. For example, the inability to hold a discourse plan in mind and monitor speech output leads to disorganized speech and thought disorder; the inability to maintain a plan for behavioral activities could lead to negative symp-

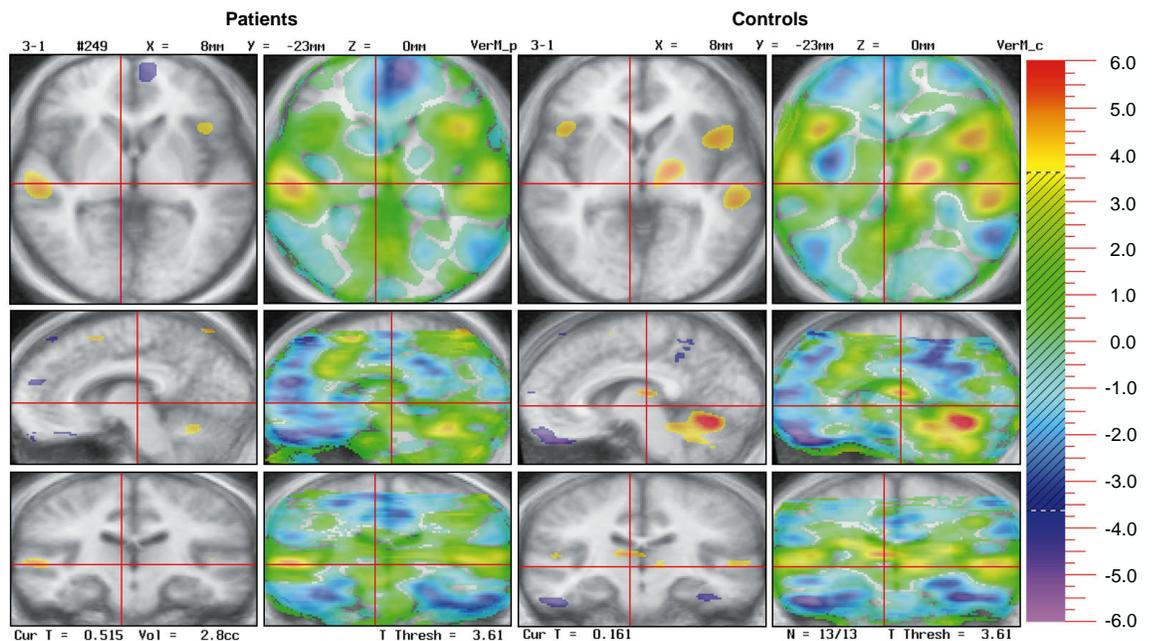
toms such as avolition or alogia; and the inability to reference a specific external or internal experience against associative memories (mediated by cortical and subcortical circuitry involving frontal/parietal/temporal regions and the thalamus) could lead to an altered consciousness of sensory experience that would be expressed as delusions or hallucinations. The model also explains the perseverative behavior observed in studies using the Wisconsin Card Sorting Test and is consistent with the compromised blood flow to the prefrontal cortex seen in these patients (22).

This model receives strong support from basic studies of cognition in nonhuman primates, which has permitted Goldman-Rakic to map the circuitry engaged by working memory tasks. Using a combination of lesion studies, tract tracing, studies of cerebral metabolism, and single-cell recordings in awake behaving monkeys, she and her colleagues have demonstrated that the dorsolateral prefrontal cortex (the principal sulcus, or Walker’s area 46) plays a key role in delayed response tasks that require the animals to hold the memory of the spatial location of a reward online. The dorsolateral prefrontal cortex is anatomically directly connected with the posterior parietal cortex, which processes and stores visuo-spatial information relayed to it from the “dorsal” visual pathway. Other dissociable prefrontal regions appear to be specialized for object

features and to have connections with inferior temporal lobe regions that code such features (23). Overall, her work supports a model that suggests a major role for prefrontal regions and their multiple distributed cortical, thalamic, and striatal connections in a fundamental cognitive function—representationally guided behavior—that permits organisms to adapt flexibly to a changing environment and to achieve temporal and spatial continuity between past experiences and present and future actions.

The applicability of this model to schizophrenia has been shown through neuropathology studies that demonstrate increased cell packing density in the prefrontal cortex in schizophrenic patients, which is consistent with a loss of neuropil (24). The pharmacologic agents used to reduce the symptoms of schizophrenia have been shown to exert therapeutic efficacy through dopaminergic blockade, with newer “atypical” neuroleptics such as clozapine having more potent effects on D1 and D4 dopamine receptors. Using iontophoresis of a selective D1 antagonist onto principal sulcus neurons in monkeys performing working memory tasks, Williams and Goldman-Rakic have shown that D1 blockade potentiates their activity, perhaps by disinhibiting their excitatory *N*-methyl-D-aspartate (NMDA) receptor input (25). Adding support to this hypothesis, the team has also used immunocytochemistry to show that a D4 receptor antibody

Fig. 2. The functional circuits used by healthy volunteers (two columns at right) and by schizophrenic patients (two columns at left) during practiced recall of complex narrative material are shown as visualized with PET. The images show the components engaged in coordinating cognitive processes through prefrontal-thalamic-cerebellar circuitry. Statistical maps (*t* maps) of the PET data, showing significantly activated regions, are superimposed on a composite MR image derived by averaging the MR scans from the individuals. Regions in red/yellow tones have higher blood flow during the memory task. Two types of statistical maps are provided. The “peak map” (left column in each pair) shows the small areas where all contiguous voxels exceed the predefined threshold for statistical significance (3.61). The “*t* map” (right column in each pair) shows the value of *t* for all voxels in the image and provides a general overview of the landscape of increases in blood flow during the task. The value of *t* is shown on the color scale at right. The controls have activity in



the bilateral frontal opercula, left thalamus, and left planum (transaxial view, top); left thalamus and cerebellum (sagittal view, middle); and bilateral thalami (right greater than left) and plana (left greater than right) (coronal view, bottom). The activity in the right frontal operculum and bilateral thalami is absent in the patients, and the cerebellar activity is much diminished. These deficits are presumed to be the basis for their cognitive dysmetria.

labeled γ -aminobutyric acid-containing inhibitory interneurons in the prefrontal cortex, hippocampus, reticular thalamus, globus pallidus, and substantia nigra; these results are also consistent with the possibility that atypical neuroleptics improve schizophrenia by disinhibiting excitatory transmission in cortical-striatal-thalamic pathways (26).

Psychobiology and neurophysiology. Using techniques originally derived from neurophysiology, Braff and colleagues have developed another complementary model. This model begins from the perspective of techniques used to measure brain electrical activity, particularly various types of evoked potentials, and hypothesizes that the core underlying deficit in schizophrenia involves information processing and attention (27). This model derives from the empirical clinical observation that patients with schizophrenia frequently complain that they are bombarded with more stimuli than they can interpret (28). Consequently, they misinterpret (that is, have delusions), confuse internal with external stimuli (hallucinations), or retreat to safety ("negative symptoms" such as alogia, anhedonia, or avolition). Early interpretations of this observation drew on the Broadbent filter theory and postulated that patients had problems with early stages in serial order processing that led to downstream effects such as psychotic or negative symptoms (29). As serial models have been supplanted by distributed models, the deficit may be better conceptualized in terms of resource allocation: Patients cannot mobilize attentional resources and allocate them to relevant tasks.

This group has explored this model using neurophysiological paradigms to examine sensory gating. Prepulse inhibition is a technique that measures gating of the startle response, which can be triggered by a bright light or loud auditory stimulus; in normals the startle response can be diminished (inhibited) if a weak prepulse stimulus is delivered a short time (milliseconds) before a strong startle-eliciting stimulus. Patients with schizophrenia have impaired prepulse inhibition, which occurs across a broad range of stimulus intensities (30). A related paradigm, gating of the P50 evoked potential by a prepulse stimulus, has also been shown by this group and others to be impaired in schizophrenic patients (31). Waldo *et al.* have also shown impairment in P50 gating in unaffected family members of patients with schizophrenia, raising the possibility that this could be used as a trait or vulnerability measure (32).

Prepulse inhibition can be readily modeled in animals and used to study the effects of medications, because all mammals display the startle response. Using lesion

methods and single-cell recordings, this group has demonstrated that cortical-striatal-pallidal-thalamic circuitry plays a key role in modulating the startle response in rats. They have found that dopamine agonists such as apomorphine can also lead to loss of sensory gating and that both typical and atypical antipsychotic medications restore prepulse inhibition in apomorphine-treated rats in a profile that correlates with their clinical efficacy, which suggests that this model could serve as a screen for neuroleptic medications, using a cognitive measure that may be related to the underlying neural defect in schizophrenia (33).

Clinical psychiatry. Our group has used the clinical presentation of schizophrenia as a point of departure, initially attempting to localize the various symptoms in brain regions through the use of structural and functional neuroimaging techniques (34). This strategy has led to a search for abnormalities in specific brain regions and theories about symptom-region relationships (such as negative symptoms in the frontal cortex or hallucinations in the superior temporal gyrus), which have been examined by a variety of investigators (35). This approach is oversimplified, however, and we are currently testing an integrated model that explains clinical symptoms as a consequence of disruptions in anatomically identified circuits that mediate a fundamental cognitive process.

Based on relatively consistent observations of abnormalities in frontal, thalamic, and cerebellar regions in schizophrenia, using both magnetic resonance imaging (MRI) and PET, we have postulated that the symptoms arise from impaired connectivity between these regions as a consequence of a neurodevelopmental defect or perhaps a series of them (36). Motor dysmetria has been observed in schizophrenia since its original description by Kraepelin, and "soft signs" of poor coordination are reported in more contemporary studies (37). More injurious, however, is the related "cognitive dysmetria," which produces "poor coordination" of mental activities (37). The word "metron" literally means "measure": A person with schizophrenia has a fundamental deficit in taking measure of time and space and in making inferences about interrelationships between self and others, or among past, present, and future. He or she cannot accurately time input and output, and therefore cannot coordinate the perception, prioritization, retrieval, and expression of experiences and ideas. This hypothesis has received strong support to date from our work with PET, which has revealed abnormalities in frontal-thalamic-cerebellar circuitry across a broad range of cognitive tasks. Figure 2 shows an image

from a PET study that illustrates these abnormalities in patients who display cognitive dysmetria (38).

The common thread. The common thread in all these observations, spun from four different starting points, is that schizophrenia reflects a disruption in a fundamental cognitive process that affects specific circuitry in the brain and that may be improved through medications that affect that circuitry. The various teams use differing terminology and somewhat different concepts—metarepresentations, representationally guided behavior, information processing/attention, cognitive dysmetria—but they convey a common theme. The cognitive dysfunction in schizophrenia is an inefficient temporal and spatial referencing of information and experience as the person attempts to determine boundaries between self and not-self and to formulate effective decisions or plans that will guide him or her through the small-scale (speaking a sentence) or large-scale (finding a job) maneuvers of daily living. This capacity is sometimes referred to as consciousness.

Using diverse technologies and techniques—PET scanning, animal models, lesion methods, single-cell recordings, evoked potentials—the investigators also converge on similar conclusions about the neuroanatomic substrates of the cognitive dysfunction. All concur that it must involve distributed circuits rather than a single specific "localization," and all suggest a key role for interrelationships among the prefrontal cortex, other interconnected cortical regions, and subcortical regions, particularly the thalamus and striatum. Animal and molecular models are being developed that are based on knowledge of this circuitry and the fundamental cognitive process, which can be applied to understanding the mechanism of drug actions and developing new medications.

Dissecting the anatomy of melancholy. Depression is the most common of the mood disorders. It is characterized by an overwhelming sense of sadness, sometimes without any obvious precipitant, accompanied by slowing of thoughts and actions (psychomotor retardation), insomnia, anorexia, alterations in normal diurnal rhythms, difficulty in concentrating, and feelings of hopelessness and guilt. Because symptoms such as insomnia or altered diurnal rhythm could be explained by a disruption in endocrine regulation, particularly involving cortisol, during the past several decades many investigators have focused on the neuroendocrinology of depression (39). There have been fewer cognitive models proposed for depression than for schizophrenia. However, most investigators agree on the core psychopathological feature of depression: a severe alteration in emotional tone that



gives a negative coloring to many aspects of the person's thoughts and behavior.

The lesion method and neuroimaging. One approach to developing a cognitive model of depression derives from the use of the lesion method. Robinson and others have examined the effects on mood of the lesions produced by stroke in humans and have explored neural mechanisms by studying the effects of comparable lesions in rats, thereby combining human and animal models. They have observed that mood disorders occur in 30 to 50% of stroke patients and that the course and severity of the mood disturbance are not related to the severity of the language, physical, or intellectual impairment produced by the stroke (40). Rather, lesion location is a crucial factor in determining the occurrence and type of mood disturbance; location explains 50 to 70% of the variance in severity (41). Depression is strongly associated with left anterior frontal lesions, whereas more posterior left frontal lesions do not produce significant mood disturbance. Further, lesions in the right hemisphere produce a manic rather than a depressive syndrome (42). Robinson hypothesizes that the changes in emotion produced by stroke are a "frontal syndrome" that varies in emotional tone according to the hemisphere injured by the lesion. Studies of secondary mania suggest that the amygdala may also be involved (43).

These changes in mood were modeled in rats by producing left or right hemisphere infarctions through ligation of the middle cerebral artery. Infarction produces a general reduction in biogenic amines, including both norepinephrine and serotonin, but it also leads to focal decreases in frontal regions (44). This neurochemical imbalance can be reversed by treatment with antidepressant medications, which affect norepinephrine and serotonin systems, indicating that an anatomical lesion can be modulated by chemical mechanisms and that the distinction between "systems" and "molecules" is arbitrary (45).

Imaging studies provide independent confirmation of many of these observations in patients suffering from clinical depression. When compared to healthy controls, patients suffering from a full depressive syndrome display decreased perfusion in left dorsolateral and bilateral medial prefrontal regions and the left posterior parietal region (46). Bench *et al.* have shown that after treatment and recovery, patients at least partially "normalize" and have relative increases in flow in frontal regions, although not in parietal regions (47). Because these frontal regions have also been implicated in "psychomotor poverty dimension" that is related to the alogia and avolition of schizo-

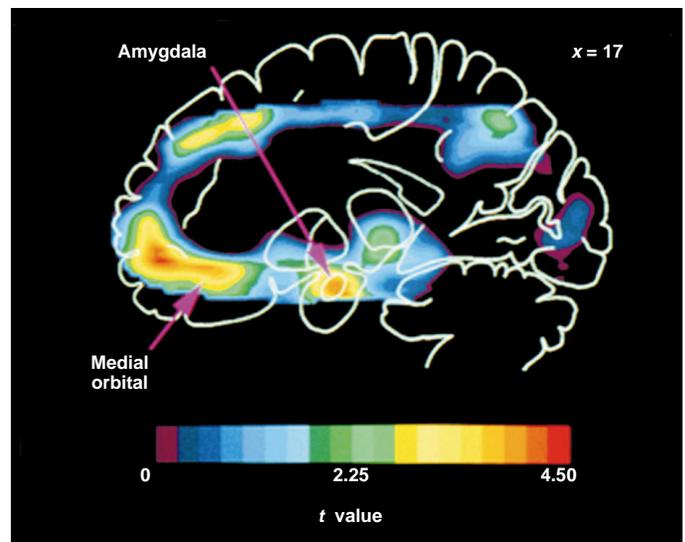
phrenia (12), and because psychomotor retardation is an important symptom of depression, these findings support the possibility that similar symptoms observed in different disease conditions may share a common neural substrate (47). Further, Parkinsonian patients, who also display psychomotor retardation, show similar decreases in frontal flow in imaging studies (48). Patients suffering from depression and schizophrenia also have similar impairments in sustained attention and similar patterns of "hypofrontality" in PET studies (49).

Classical conditioning in animals and cognitive models of mood. Classical conditioning, once treated as a "black box" by behaviorists, is now being studied by neuroscientists who are identifying how the brain learns to link a response to a stimulus. LeDoux and Davis have developed a model that uses conditioning to explore the relations among emotion, memory, and the brain (50), and their models provide additional insights into the neural mechanisms of mood disorders. Combining "fear conditioning" in rats with the lesion method and neurophysiological techniques, they have explored the neuroanatomy of emotion. This work indicates that exposure to an aversive experience lays down powerful memory traces that are difficult to extinguish, stored in a circuit in which the amygdala plays a central role, but that has thalamic, hippocampal, and cortical components. There appear to be two parallel tracks for processing emotional stimuli; a

fast track engages subcortical regions (the thalamus and amygdala) and permits rapid but relatively crude responses, whereas the slower track permits more refined identification of features and formulation of plans and includes the prefrontal cortex (51). The hippocampus also plays a role by assessing contextual cues (52). The molecular mechanism for the formation of emotional memories appears to involve glutamatergic activation of NMDA receptors (53).

This work provides a cognitive model for examining the neural substrates of human illnesses such as depression or anxiety. Most cognitive theories of depression posit that it results from the effects of aversive experiences that are stored as memory traces and that, if reactivated by appropriate cues, will produce the characteristic negative emotional tone of depression. "Learned helplessness," the experience of being repeatedly exposed to physical or mental trauma from which there is no opportunity to escape, is an example of experimentally induced negative memories that lead to depressive syndromes in both animals and humans (54). The impact of the aversive memories may be accessible to consciousness, giving the person an insight as to why depression is triggered at a given time or in a given place, but the impact may also be unconscious, triggering a depression that appears to be "endogenous." Because the amygdala has strong input to the hypothalamus, the neuroendocrinologic aspects of depression are also explicable: Increasing activity in the amygdala will also trigger adrenocorticotrop-

Fig. 3. Areas of increased blood flow in the left amygdala and the medial orbital cortex of people with familial, major depressive disorder are shown. The sagittal image slice shown is from an image of *t* values, produced by a voxel-by-voxel computation of the *t* statistic, and indicates areas where flow is increased in depressives relative to controls. The statistical significance of abnormal activity in these regions has been confirmed with both regional blood flow and glucose metabolic image data from multiple, independent samples of individuals. The area of increased flow in the left prefrontal cortex extends from the medial orbital cortex to also involve areas of the lateral orbital, ventrolateral, and ventromedial portions of the prefrontal cortex. The *x* coordinate locates the image section in millimeters to the left of the midline. Anterior is toward the reader's left. [Reprinted from J. L. Price, S. T. Carmichael, W. C. Drevets, *Progr. Brain Res.* **107**, 523 (1996), with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, Netherlands]



ic hormone release in the hypothalamus, leading in turn to corticosteroid release.

Recent neuroimaging studies in humans have provided support for a neural model of depression that involves a thalamic-amygdala-frontal loop similar to that observed in conditioning studies in animals. Drevets *et al.* used PET to compare active and remitted depressives with a strong family history of mood disorder with normal volunteers in order to explore state versus trait differences (55). Hypothetically, negative emotional memories are a "trait" that is always present; their activation leads to the clinical state known as depressive disorder, whereas their suppression or modulation (either by pharmacotherapy or psychotherapy) leads to remission of symptoms. Drevets *et al.* found increased activity in the left amygdala in both active and remitted depressives, but only the active depressives had increased left frontal activity, which suggests that the prefrontal-amygdala pathway may be activated during the depressed state. This is illustrated in Fig. 3. In addition to the frontal regions and amygdala, increased activity was also observed in the thalamus, putamen, and hippocampus.

The basic structure. Studies of mood disorders concur on the general nature of the underlying cognitive process of depression: It is a pathological alteration in emotion. Studies using the lesion method, behavioral conditioning, and PET have begun to dissect the anatomy of emotion: Portions of the prefrontal cortex (probably anterior and inferior) and the amygdala are key nodes in the circuit, and the parietal lobes and other regions may play a role as well. Although the anatomical circuits clearly implicate frontal regions and related circuitry, the precise cognitive mechanisms are currently an area of active investigation. The data from conditioning studies and PET studies of active versus remitted patients suggest an interesting hypothesis: Memories of past pain are retained in regions such as the amygdala or parietal cortex and may lie dormant ("the trait"), predisposing an individual to developing a clinical depression ("the state") if additional factors arise. A key point in this work, however, is that the anatomy of melancholy can be modified by both psychological and chemical/molecular experiences. The depressed state can often be reversed through treatment with drugs that affect the biogenic amine systems of the brain, but it can also be treated with cognitive therapies that attempt to reverse "negative sets," and combination therapies are perhaps the most effective of all. Depression may arise as a consequence of the plastic response of mind/brain to experience, and it may also remit because of

either pharmacologic or psychotherapeutic manipulations of brain plasticity.

Summary and Conclusion

Examples of work applying diverse techniques of cognitive neuroscience to the study of depression and schizophrenia indicate that increasingly sophisticated strategies and conceptualizations are emerging as powerful new technologies are being applied. Focal regions have been replaced by circuits and static changes by plasticity and molecular mechanisms. The power of models is enhanced by efforts to design experiments that can be used in nonhuman species, in order to obtain *in vivo* measures that will illuminate mechanisms. The power of neuroimaging is also permitting *in vivo* measures of circuits and mechanisms in the human brain. These advances have created an era in which a scientific psychopathology that links mind and brain has become a reality.

REFERENCES AND NOTES

1. J. R. Searle, *The Rediscovery of the Mind* (MIT Press, Cambridge, MA, 1992); D. C. Dennett, *Kinds of Minds: Toward an Understanding of Consciousness* (Basic Books, New York, 1996).
2. M. M. Merzenich and K. Sameshima, *Curr. Opin. Neurobiol.* **3**, 187 (1993); W. Singer, *Science* **270**, 758 (1995).
3. M. I. Posner and S. E. Peterson, *Annu. Rev. Neurosci.* **13**, 25 (1990); E. Tulving, *Elements of Episodic Memory* (Oxford Univ. Press, New York and Oxford, 1983); J. L. McGaugh, N. M. Weinberger, G. Lynch, Eds., *Brain Organization and Memory: Cells, Systems and Circuits* (Oxford Univ. Press, New York, 1990); L. G. Ungerleider, *Science* **270**, 769 (1995).
4. L. R. Squire and S. Zola-Morgan, *The Physiological Basis of Memory* (Academic Press, New York, 1983); P. Goldman-Rakic, *Trends Neurosci.* **7**, 425 (1984); B. Milner and M. Petrides, *ibid.*, p. 403; A. L. Benton, *Contributions to Clinical Neuropsychology* (Aldine, Chicago, 1969); N. Geschwind and A. M. Galaburda, Eds., *Cerebral Dominance: The Biological Foundations* (Harvard Univ. Press, Cambridge, MA, 1984).
5. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)* (American Psychiatric Press, Washington, DC, 1994); T. A. Widiger *et al.*, Eds., *DSM-IV Sourcebook, Volume 1* (American Psychiatric Press, Washington, DC, 1994).
6. E. Robins and S. B. Guze, *Am. J. Psychiatry* **126**, 983 (1970).
7. N. C. Andreasen, *Schizophr. Bull.* **13**, 9 (1987).
8. J. P. Feighner *et al.*, *Arch. Gen. Psychiatry* **26**, 57 (1972).
9. J. M. Hall *et al.*, *Science* **250**, 1684 (1990); A. D. Roses, *Annu. Rev. Med.* **47**, 387 (1996).
10. E. Bleuler, *Dementia Praecox of the Group of Schizophrenias (1911)*, J. Zinkin, Transl. (International Universities Press, New York, 1950).
11. C. D. Frith, *The Cognitive Neuropsychology of Schizophrenia* (Erlbaum, East Sussex, United Kingdom, 1992).
12. P. F. Liddle *et al.*, *Br. J. Psychiatry* **160**, 179 (1992); C. D. Frith, K. Friston, P. F. Liddle, R. S. Frackowiak, *Proc. R. Soc. London Ser. B* **244**, 241 (1991).
13. C. Frith, *Lancet* **346**, 615 (1995).
14. P. C. Fletcher, C. D. Frith, P. M. Grasby, K. J. Friston, R. J. Dolan, *J. Neurosci.* **16**, 7055 (1996).
15. P. K. McGuire and C. D. Frith, *Psychol. Med.* **26**, 663 (1996).

16. P. K. McGuire *et al.*, *ibid.*, p. 29.
17. P. K. McGuire *et al.*, *Br. J. Psychiatry* **169**, 148 (1996).
18. D. A. Silbersweig *et al.*, *Nature* **378**, 176 (1995).
19. P. C. Fletcher *et al.*, *Cognition* **57**, 109 (1995).
20. P. S. Goldman-Rakic, *J. Neuropsychiatry Clin. Neurosci.* **6**, 348 (1994).
21. _____, in *Handbook of Physiology* F. Plum and V. Mountcastle, Eds. (American Physiological Society, Bethesda, MD, 1987), pp. 373-417.
22. D. R. Weinberger, K. F. Berman, R. F. Zec, *Arch. Gen. Psychiatry* **43**, 114 (1986).
23. P. S. Goldman-Rakic, *Trends Neurosci.* **7**, 419 (1984); in *Neurobiology of Neocortex*, P. Rakic and W. Singer, Eds. (Wiley, New York, 1988); *Neuron* **14**, 477 (1995); H. R. Friedman, J. D. Janas, P. S. Goldman-Rakic, *J. Cogn. Neurosci.* **2**, 18 (1990); P. S. Goldman-Rakic and L. J. Porrino, *J. Comp. Neurol.* **242**, 535 (1985).
24. L. D. Selemon, G. Rajkowska, P. S. Goldman-Rakic, *Arch. Gen. Psychiatry* **52**, 805 (1995).
25. G. V. Williams and P. S. Goldman-Rakic, *Nature* **376**, 572 (1995).
26. L. Mrzljak *et al.*, *ibid.* **381**, 345 (1996).
27. D. L. Braff, *Schizophr. Bull.* **19**, 233 (1993).
28. A. McGhie and J. Chapman, *Br. J. Med. Psychol.* **34**, 103 (1961).
29. D. E. Broadbent, *Perception and Communication* (Pergamon, London, 1958).
30. D. L. Braff, C. Grillon, M. A. Geyer, *Arch. Gen. Psychiatry* **49**, 206 (1992).
31. L. L. Judd, L. A. McAdams, B. Budnick, D. L. Braff, *Am. J. Psychiatry* **149**, 448 (1992).
32. M. Waldo *et al.*, *Schizophr. Res.* **12**, 93 (1994).
33. N. R. Swerdlow, D. L. Braff, N. Taaid, M. A. Geyer, *Arch. Gen. Psychiatry* **51**, 139 (1994).
34. N. C. Andreasen, *Can Schizophrenia be Localized in the Brain?* (American Psychiatric Press, Washington, DC, 1986); L. E. DeLisi *et al.*, *Biol. Psychiatry* **29**, 159 (1991); B. Bogerts *et al.*, *ibid.* **33**, 236 (1993); R. B. Zipursky *et al.*, *ibid.* **35**, 501 (1994); R. W. Buchanan *et al.*, *Am. J. Psychiatry* **150**, 59 (1993); D. R. Weinberger, K. F. Berman, R. Suddath, E. F. Torrey, *ibid.* **149**, 890 (1992); B. Turetsky *et al.*, *Arch. Gen. Psychiatry* **52**, 1061 (1995).
35. E. Barta, G. D. Pearlson, R. E. Powers, S. S. Richards, L. E. Tune, *Am. J. Psychiatry* **147**, 1457 (1990); M. E. Shenton *et al.*, *N. Engl. J. Med.* **327**, 604 (1992); N. C. Andreasen *et al.*, *Arch. Gen. Psychiatry* **49**, 943 (1992); D. H. Ingvar and G. Franzen, *Lancet* **ii**, 1484 (1974); M. S. Buchsbaum *et al.*, *Arch. Gen. Psychiatry* **39**, 251 (1982).
36. N. C. Andreasen, *Neuron* **16**, 697 (1996); _____ *et al.*, *J. Am. Med. Assoc.* **272**, 1763 (1994); N. C. Andreasen *et al.*, *Science* **266**, 294 (1994); V. W. Swayze II, N. C. Andreasen, J. C. Ehrhardt, R. Alliger, G. Cohen, *Arch. Neurol.* **47**, 805 (1990); P. Nopoulos *et al.*, *Biol. Psychiatry*, in press.
37. E. Kraepelin, R. M. Barclay, G. M. Robertson, *Dementia Praecox and Paraphrenia* (E&S Livingstone, Edinburgh, United Kingdom, 1919); S. Gupta, W. C. Hubbard, S. Arndt, M. A. Flaum, N. C. Andreasen, *Schizophr. Res.* **11**, 195 (1994).
38. N. C. Andreasen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9985 (1996).
39. F. Holsboer, in *Handbook of Affective Disorders*, E. S. Paykel, Ed. (Churchill Livingstone, Edinburgh, United Kingdom, 1992).
40. S. E. Starkstein and R. E. Robinson, *Br. J. Psychiatry* **154**, 170 (1989).
41. R. G. Robinson, K. L. Kubos, L. B. Starr, K. Rao, T. R. Price, *Brain* **107**, 81 (1984).
42. S. E. Starkstein, G. D. Pearlson, J. Boston, R. G. Robinson, *Arch. Neurol.* **44**, 1069 (1987).
43. R. G. Robinson, J. D. Boston, S. E. Starkstein, T. R. Price, *Am. J. Psychiatry* **145**, 172 (1988).
44. R. G. Robinson, *Science* **205**, 707 (1979).
45. J. R. Lipsey, R. G. Robinson, G. D. Pearlson, K. Rao, T. R. Price, *Lancet* **i**, 297 (1984).
46. L. R. Baxter Jr. *et al.*, *Arch. Gen. Psychiatry* **49**, 681 (1989); J.-L. Martinot *et al.*, *Am. J. Psychiatry* **147**, 1313 (1990).
47. C. J. Bench, R. S. J. Frackowiak, R. J. Dolan, *Psychol. Med.* **25**, 247 (1995); C. J. Bench *et al.*, *ibid.* **22**, 607 (1992).



48. H. S. Mayberg *et al.*, *Ann. Neurol.* **28**, 57 (1990).
 49. R. M. Cohen *et al.*, *Neuropsychopharmacology* **2**, 241 (1989).
 50. J. E. LeDoux, *Sci. Am.* **6**, 50 (June 1994); M. Davis, *Annu. Rev. Neurosci.* **15**, 353 (1992).
 51. J. E. LeDoux, *Curr. Opin. Neurobiol.* **2**, 191 (1992); L. M. Romanski and J. E. LeDoux, *J. Neurosci.* **12**, 4501 (1992); J. L. Armony, J. D. Cohen, D. Servan-Schreiber, J. E. LeDoux, *Behav. Neurosci.* **109**, 246 (1995).
 52. K. P. Corodimas and J. E. LeDoux, *Behav. Neurosci.* **109**, 613 (1995).
 53. M. J. D. Miserendino, C. B. Sananes, K. R. Melia, M. Davis, *Nature* **345**, 716 (1990); C. Farb *et al.*, *Brain Res.* **593**, 145 (1992); M. Davis, D. Rainne, M. Cassell, *Trends Neurosci.* **17**, 208 (1994).
 54. M. E. P. Seligman, *J. Abnorm. Psychol.* **74**, 1 (1976); F. Schneider *et al.*, *Am. J. Psychiatry* **153**, 206 (1996).
 55. W. C. Drevets *et al.*, *J. Neuroscience* **12**, 3628 (1992).
 56. Supported in part by National Institute of Mental Health grants MH31593, MH40856, and MH-CRC43271; by a Research Scientist Award, MH00625; and by an Established Investigator Award from the National Association for Research in Schizophrenia and Affective Disorders.

A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague*

The capacity to predict future events permits a creature to detect, model, and manipulate the causal structure of its interactions with its environment. Behavioral experiments suggest that learning is driven by changes in the expectations about future salient events such as rewards and punishments. Physiological work has recently complemented these studies by identifying dopaminergic neurons in the primate whose fluctuating output apparently signals changes or errors in the predictions of future salient and rewarding events. Taken together, these findings can be understood through quantitative theories of adaptive optimizing control.

An adaptive organism must be able to predict future events such as the presence of mates, food, and danger. For any creature, the features of its niche strongly constrain the time scales for prediction that are likely to be useful for its survival. Predictions give an animal time to prepare behavioral reactions and can be used to improve the choices an animal makes in the future. This anticipatory capacity is crucial for deciding between alternative courses of action because some choices may lead to food whereas others may result in injury or loss of resources.

Experiments show that animals can predict many different aspects of their environments, including complex properties such as the spatial locations and physical characteristics of stimuli (1). One simple, yet useful prediction that animals make is the probable time and magnitude of future rewarding events. "Reward" is an operational concept for describing the positive value that a creature ascribes to an object, a behavioral act,

or an internal physical state. The function of reward can be described according to the behavior elicited (2). For example, appetitive or rewarding stimuli induce approach behavior that permits an animal to consume. Rewards may also play the role of positive reinforcers where they increase the frequency of behavioral reactions during learning and maintain well-established appetitive behaviors after learning. The reward value associated with a stimulus is not a static, intrinsic property of the stimulus. Animals can assign different appetitive values to a stimulus as a function of their internal states at the time the stimulus is encountered and as a function of their experience with the stimulus.

One clear connection between reward and prediction derives from a wide variety of conditioning experiments (1). In these experiments, arbitrary stimuli with no intrinsic reward value will function as rewarding stimuli after being repeatedly associated in time with rewarding objects—these objects are one form of unconditioned stimulus (US). After such associations develop, the neutral stimuli are called conditioned stimuli (CS). In the descriptions that follow, we call the appetitive CS the sensory cue and the US the reward. It should be kept in mind, however, that learning that depends on CS-US pairing takes many different forms and is not always dependent on reward (for example, learning associated

with aversive stimuli). In standard conditioning paradigms, the sensory cue must consistently precede the reward in order for an association to develop. After conditioning, the animal's behavior indicates that the sensory cue induces a prediction about the likely time and magnitude of the reward and tends to elicit approach behavior. It appears that this form of learning is associated with a transfer of an appetitive or approach-eliciting component of the reward back to the sensory cue.

Some theories of reward-dependent learning suggest that learning is driven by the unpredictability of the reward by the sensory cue (3, 4). One of the main ideas is that no further learning takes place when the reward is entirely predicted by a sensory cue (or cues). For example, if presentation of a light is consistently followed by food, a rat will learn that the light predicts the future arrival of food. If, after such training, the light is paired with a sound and this pair is consistently followed by food, then something unusual happens—the rat's behavior indicates that the light continues to predict food, but the sound predicts nothing. This phenomenon is called "blocking." The prediction-based explanation is that the light fully predicts the food that arrives and the presence of the sound adds no new predictive (useful) information; therefore, no association developed to the sound (5). It appears therefore that learning is driven by deviations or "errors" between the predicted time and amount of rewards and their actual experienced times and magnitudes [but see (4)].

Engineered systems that are designed to optimize their actions in complex environments face the same challenges as animals, except that the equivalent of rewards and punishments are determined by design goals. One established method by which artificial systems can learn to predict is called the temporal difference (TD) algorithm (6). This algorithm was originally inspired by behavioral data on how animals actually learn predictions (7). Real-world applications of TD models abound. The predictions learned by TD methods can also be used to implement a technique called dynamic programming, which specifies how a system can come to choose appropriate actions. In this article, we review how these computational methods provide an interpretation of the activity of dopamine neurons thought to mediate reward-processing and reward-dependent learning. The connection between the computational theory and the experimental results is striking and provides a quantitative framework for future experiments and theories on the computational roles of ascending monoaminergic systems (8–13).

W. Schultz is at the Institute of Physiology, University of Fribourg, CH-1700 Fribourg, Switzerland. E-mail: Wolfram.Schultz@unifr.ch P. Dayan is in the Department of Brain and Cognitive Sciences, Center for Biological and Computational Learning, E-25 MIT, Cambridge, MA 02139, USA. E-mail: dayan@ai.mit.edu P. R. Montague is in the Division of Neuroscience, Center for Theoretical Neuroscience, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA. E-mail: read@bcm.tmc.edu

*To whom correspondence should be addressed.