13 Stabilizing Mental Disorders: Prospects and Problems

Jacqueline Sullivan

A primary focus of the debates in philosophy of psychiatry addressed in each of the chapters in this volume is whether mental disorders are natural kinds. The question subdivides into several interrelated questions: Are mental disorders real and stable regularities in nature that exist independent of our systems of classifying them? Do the sets of necessary and sufficient conditions that constitute the categories of mental disorders put forward in the *Diagnostic and Statistical Manual of Mental Disorders* and in the *International Classification of Diseases* track these regularities? Are those groups of phenomena individuated by the categories suitable for discovering their causes and identifying viable targets for therapeutic intervention?

The vast majority of philosophers of psychiatry are realists about mental disorders. The consensus, however, is that current mental disorder categories do not pick out stable regularities in nature that are subject to the same causal-mechanical explanations. (See, e.g., Craver 2009; Kendler, Zachar, and Craver 2011; Haslam 2000, 2002, and this volume; Insel 2013; see also the chapters in this volume by Kincaid and Murphy.) Yet if the categories do not track real divisions in nature—if research into mental disorders begins with indefinite and poorly circumscribed explanatory targets—it is likely that the projects of identifying their causes and developing successful therapeutic interventions to treat them will also fail. That scientific explanation requires well-delineated explanatory targets, and mental disorders do not seem to qualify, is one of the primary reasons why philosophers of psychiatry have been reluctant to abandon the natural kinds ideal for psychiatric classification and why debates about whether or not mental disorders are natural kinds persist in the philosophical literature.

Some of the chapters in this volume (those by Kincaid, Horwitz, Murphy, and Ross) focus on how to revise current categories of mental disorders so that the disorders they individuate correspond to bona fide regularities in

nature or something close enough. In this chapter, I take a slightly different approach. Specifically, I consider the instability to which psychiatric kinds may be subject when they become explanatory targets of areas of science that are not "mature" (see, e.g., Hacking 1988, 1992) and are in the early stages of discovering the mechanisms of cognitive phenomena (see, e.g., Bechtel and Richardson 1993; Bechtel 2008; Craver 2007). I focus primarily on two such areas of science that have been independently involved in the investigation of the mechanisms of mental disorders: cognitive neuroscience (which has as its task the localization of cognitive functions in the brain) and cognitive neurobiology (which aims to discover the cellular and molecular mechanisms of learning and memory).

There is a growing consensus among research scientists that mental disorders are simply disorders of cognition. (See, e.g., Carter et al. 2009; Insel 2013; Nuechterlein et al. 2012.) This has led to the emergence of intra-disciplinary and inter-disciplinary research initiatives to identify the cognitive functions and the underlying synaptic, cellular, and molecular mechanisms that are disrupted in mental disorders, the ultimate aim being the development of effective therapeutic interventions. In this chapter, I evaluate one such research initiative: the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS) initiative. I use this case study as a basis to show that such research does not begin with "stable phenomena" (see, e.g., Hacking 1988, 1992) that track mechanisms; neither the mental disorders nor the cognitive functions under study qualify as such. The reason for this instability is that the methods used to individuate these explanatory targets are not standardized either within or across disciplinary boundaries. Thus, for inter-disciplinary interactions to be effective, specific measures must be taken across different contexts of experimentation to ensure the stability of the phenomena under study. What is interesting about the CNTRICS initiative is that investigators have sought to impose intra-disciplinary and inter-disciplinary "strategies of stabilization" (see, e.g., Hacking 1983, 1988, 1992) to operationally fix the cognitive functions that are disrupted in schizophrenia across investigators and levels of analysis. I assess the potential for these strategies to succeed at the goal of stabilizing mental disorders as scientific kinds.

Mental Disorders and Stability

Ian Hacking has discussed the stability of scientific kinds in two separate contexts. First, he has used historical case studies to demonstrate that

categories of mental disorders and the phenomena to which they refer are unstable. (See, e.g., Hacking 1995a,b, 2007.) Such instability originates in the context of psychiatric diagnosis when an individual is informed that he or she has a mental disorder, and that he or she may thus be classified as a specific kind of person. Second, he has shown that the stability of the kinds under study in laboratory sciences is very much tied to how established a particular laboratory science is. (See, e.g., Hacking 1983, 1988, 1989, 1992.) In mature areas of science we often encounter stable kinds of phenomena under study, whereas in immature laboratory sciences the kinds have a tendency to oscillate wildly. The source of stability or instability in this case arises in the context of the laboratory and has to do in part with investigators experimentally harnessing a phenomenon. Although in this chapter I am primarily interested in this latter kind of stability, I want to say something about Hacking's worry that psychiatric diagnosis may be a source of the instability of psychiatric kinds because it has featured prominently in philosophical debates about whether mental disorders are natural kinds. (See, e.g., Bogen 1988; Cooper 2004; Khalidi 2010; Tekin, this volume.)

Hacking characterizes mental disorders as paradigmatic examples of kinds that lack stability because they are subject to what he dubs "looping effects." (See, e.g., Hacking 1995a,b, 2007.) Specifically, classifying a human subject as a kind of thing (e.g., a "manic depressive") or diagnosing that subject with a mental disorder may prompt changes in that individual such that the criteria that constitute the mental-disorder category are no longer applicable and require revision. Such revisions, if they are made at all, are only stopgap measures; they do not guarantee the stability of psychiatric kinds since future looping effects are always possible. Categories of mental disorders, thus, can never be stable, because the kinds they pick out are "moving targets." (See, e.g., Hacking 2007.)

Hacking's "looping effects" argument is convincing in part because the effects are so plausible—human beings do adjust their behaviors in response to being categorized. Furthermore, in the cases of multiple personality disorder (1995b) and dissociative fugue (1998) Hacking puts forward two convincing examples of looping effects in action that make it seem plausible that other mental disorders may be subject to similar types of effects. However, if psychiatric kinds are indeed unstable in the way Hacking suggests, this has negative implications not only for psychiatry and psychiatric diagnosis but also for any area of science that is in the business of investigating the mechanisms of mental disorders and developing successful therapeutic interventions for treating them, for truly explaining a

phenomenon seems to require a stable and fixed explanatory target. (See, e.g., Bechtel 2008; Bechtel and Richardson 1993; Craver 2007; Sullivan 2010.) If Hacking is correct and categories of mental disorders have moveable rather than fixed referents, then we will not be able to explain and treat mental disorders.

The response in the philosophical literature to Hacking's claim about mental disorders' being subject to looping effects has been mixed. (See, e.g., the chapter by Tekin in this volume.) However, philosophers of psychiatry have, for the most part, dismissed looping effects as real obstacles to discovering the causes of mental disorders, for two reasons. First, although it is plausible that individuals may change in detectable (and undetectable) ways in response to a diagnosis of a mental disorder, no evidence exists to establish that the process actually happens for all categories of mental disorders and all individuals or to what extent. This leaves open the possibility that at least some mental disorders are kinds that track relatively stable regularities in nature that do not change radically in response to classification. (See, e.g., the chapters by Haslam, Kincaid, Murphy, Ross, and Zachar).

Another strategy for responding to Hacking's worry about looping effects is to demonstrate that a lot of important causal discoveries are made in science in the absence of stable or "natural" kinds. For example, Kincaid (2008, p. 373) argues that we have "piecemeal causal explanations of cancer," and that we have learned how to intervene in various forms of it and treat patients with it even though what "constitutes a cancerous cell" cannot be specified "in terms of necessary and sufficient conditions." He claims that the kinds of causal explanations we have of mental disorders are similarly fragmentary. He points to the example of depression, for which we have "piecemeal causal explanations" (e.g., low levels of the neurotransmitter serotonin at serotonergic synapses in the brain) and treatments (e.g., selective serotonin reuptake inhibitors) even though it is not a natural kind. Though Kincaid does not regard such piecemeal causal explanations of mental disorders as ultimately satisfactory, his primary aim is to demonstrate that science often proceeds quite successfully in the absence of natural kinds and well-developed theories of those kinds.

Such piecemeal explanations of mental disorders arise in part from the fact that mental disorders are complex phenomena: the biological systems that exhibit them are ontologically complex, consisting of physical parts and processes that span multiple levels of organization, from genes to neurons to behavior. Because each constitutive level may be probed in

order to identify the causes of a mental disorder, the sciences directed at investigating these causes are many and diverse, spanning multiple levels of analysis from molecular genetics to behavior and incorporating a wide range of techniques—from functional imaging studies on human subjects to pharmacological intervention techniques in animal models. Thus, when we look at the scientific literature on mental disorders we encounter what William Wimsatt says we find with respect to the scientific study of complex phenomena more generally: "a plurality of incompletely articulated and partially contradictory, partially supplementary theories and models," each taken in isolation having "individually, [an] impoverished view[] of [its] objects" (2007, p. 180). Such pluralism is an important contributing factor to the instability of mental disorders as kinds of phenomena insofar as each area of science has different assumptions about the best way to operationalize mental disorders (e.g., which measurement techniques to use), where to look for the mechanisms (i.e., where in the brain or world and at what level of organization), how to look for them (i.e., different methodological strategies), and where to intervene so as to determine causal relationships (e.g., which neurotransmitter system, which receptors).

However, inter-disciplinary pluralism is not the only obstacle to stabilizing mental disorders as scientific kinds. As I have argued previously (e.g., in Sullivan 2009), intra-disciplinary methodological pluralism is also an obstacle to relating explanations of phenomena across multiple levels of analysis. Often researchers working in the same area of science but in different laboratories vary the methods used for studying a phenomenon, which leaves open the possibility that they are not all investigating the same phenomenon. (See Sullivan 2009.) This kind of local inter-disciplinary pluralism is characteristic of what Hacking refers to as "immature laboratory sciences" (1992, p. 57). The phenomena under study in such sciences are unstable "in part because [they] are produced by fundamentally different techniques, and different theories answer to different phenomena that are only loosely connected" (ibid.).

Whereas in mature areas of science the "theories and laboratory equipment [have] evolve[d] in such a way that they match each other and are self-vindicating" (Hacking 1992, p. 56), as I show later in the chapter, in those laboratory sciences (e.g., cognitive neuroscience and cognitive neurobiology) that have come to direct their attention to investigating the mechanisms of mental disorders, the requisite symbiosis between what phenomena investigators take their investigative strategies to measure and what phenomena are actually being measured is absent. These areas of

science fail to be "self-vindicating" on Hacking's understanding of the term, insofar as there is a lack of coordination among investigators with respect to how to produce, measure, and detect what they refer to as the same phenomenon. Investigators in both areas of neuroscience may be described as sharing a set of background assumptions, basic methods, investigative strategies, and explanatory goals (even a "Kuhnian paradigm"), but both areas afford investigators the flexibility to modify specific aspects of standard tasks or experimental paradigms in ways that affect, for example, what cognitive functions, areas of the brain, and cellular and molecular activities are involved in a given experiment. Such flexibility actually promotes the development of a plurality of explanations of mental disorder phenomena rather than the discovery of single unified explanations for such phenomena.

Scientists do not regard such piecemeal causal explanations as an adequate stopping point for research into the causes of mental disorders. For example, in a recent blog post the director of the US National Institutes for Mental Health, Thomas Insel, encouraged research scientists to develop "a new nosology of mental disorders" that understands them as "biological disorders involving brain circuits that implicate specific domains of cognition, emotion, or behavior," the study of which requires that "each level of analysis . . . be understood across a dimension of function" (Insel 2013). However, achieving this aim requires that the phenomena in question be stable or that investigators reach consensus that the term used to refer to a particular function or a particular mental disorder had the "same" referent.

However, such translatability doesn't simply emerge. Scientists have to impose both inter-disciplinary and intra-disciplinary standards in order to achieve it. Recent work to discover the mechanisms that mediate cognitive dysfunction in schizophrenia indicates that scientists must actively structure practice in such a way so as to ensure the stability of their explanatory targets, because the goals they aim to achieve are not thought to be attainable by any other means. Even though ensuring the stability of scientific kinds appears attractive with respect to specific explanatory and therapeutic goals, it may not be attainable, and if realized it may lead to undesirable consequences.

Schizophrenia Research

Schizophrenia is classified as a psychotic disorder that, on "the narrowest definition," involves "distortions or exaggerations in inferential thinking"

(i.e., delusions) and "in perception" (i.e., hallucinations) (American Psychiatric Association 1994, p. 274). On the *DSM*-IV definition, a set of necessary and sufficient conditions must be satisfied in order for a person to be diagnosed as having schizophrenia. First, the individual must exhibit at least two of the following characteristic symptoms, which must be present for at least a month: "delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms, i.e., affective flattening, alogia, or avolition" (ibid., p. 285). The symptoms must result in social and/or occupational dysfunction, and the "continuous signs of the disturbance" must "persist for at least 6 months" (ibid., p. 285). In addition, the diagnosis must rule out that the symptoms are due to substance use or a general medical condition and must exclude that the person suffers from Schizoaffective Disorder, Mood Disorder, or Pervasive Developmental Disorder. Insofar as many persons with schizophrenia have symptoms that are differentially manifest, some symptoms (e.g., paranoia or disorganized thoughts) being more prominent than others, the disorder has five identified subtypes to accommodate these differences: paranoid, disorganized, catatonic, undifferentiated, and residual.

The primary symptoms of schizophrenia are characterized as "cognitive and emotional dysfunctions" and are divided into two classes: positive symptoms, so called because they involve "an excess or distortion of normal functions," and negative symptoms, which "appear to reflect a diminution or loss of normal function" (American Psychiatric Association 1994, p. 274). Although anti-psychotics and neuroleptics, when taken regularly, eliminate or mask the delusions and the hallucinations associated with schizophrenia, no reliable treatments exist for the negative symptoms (e.g., disorganization in speech and behavior and affective flattening). Some experts regard these negative symptoms, considered independent of side effects from anti-psychotics, as constituting a set of "core cognitive deficits," which they regard as primary obstacles to persons diagnosed with schizophrenia maintaining steady employment and meaningful interpersonal relationships (Green, Lee, and Kern 2009, p. 158). Such failures to "achieve adequate community functioning, including finding a job, forming a network of friends or living independently" (ibid., p. 160) are taken to be responsible for "up to 40% of the excess premature mortality" in schizophrenics that "may be attributed to suicide and unnatural deaths" (Hor and Taylor 2010, p. 81).

The "cognitive deficits [that] have been demonstrated at multiple levels in schizophrenia" are said to "range from early sensory processing deficits" in both auditory and visual processing to "higher order

information-processing deficits" in "attention, executive function, working and episodic memory, and affective processing" (Belger and Barch 2009, p. 303). Cognitive neuroscientists have been careful to differentiate these cognitive deficits from the negative clinical symptoms identified in the *DSM* as indicative of schizophrenia. In fact, research scientists on the whole have been critical of the lack of validity of the categories of mental disorders identified in the *DSM*. Cameron Carter and colleagues express this sentiment best in claiming that, although the *DSM* and similar manuals "segregate mental disorders into distinct categories,"

> in actuality these disorders may reflect a complex combination of disturbances in more fundamental processes, or dimensions of function, that do not necessarily align with currently identified categories of disorder. For example, schizophrenia and depression may ultimately be better understood as particular changes in the functioning of underlying cognitive and emotional systems (such as executive control and reinforcement learning) with different combinations of deficits in these underlying systems producing different behavior deficits that are categorized clinically as two distinct illnesses. If this is so, then it suggests that a more powerful way of characterizing and understanding these illnesses is to draw on the conceptual and experimental framework provided by cognitive psychology by developing behavioral paradigms that can sensitively and specifically measure the critical functions of interest. This also promises a more direct path toward relating clinical disturbance to underlying neurobiological pathology, insofar as fundamental cognitive processes are more likely to map onto specific identifiable neurobiological mechanisms. (Carter et al. 2009, p. 169)

Carter and other cognitive neuroscientists are advocating for an inter-disciplinary initiative to revise current categories of mental disorders, with the ultimate aim of "understanding the neurobiological underpinning of . . . cognitive deficits" in mental disorders like schizophrenia by "linking [them] to selected cortical and subcortical neural circuits" (Belger and Barch 2009, p. 303). Such discoveries would then pave the way for cognitive neurobiologists who study the synapses, cells, and molecules that constitute these neural circuits to investigate the cellular and molecular mechanisms that underlie cognitive dysfunctions in schizophrenia in animal models and to test pharmacological interventions. This might lead to translational work from animal models to clinical trials to test experimental drugs in human schizophrenics.

I think the vast majority of researchers who are interested in understanding, explaining, and treating mental disorders agree that the current crisis in mental health will be resolved only by means of inter-disciplinary initiatives and that we must move beyond what everyone regards as the

inadequacy of current categories of mental disorders for tracking mental-disorder mechanisms. However, initiating and sustaining collaborative and coordinated efforts among investigators working in two radically different areas of neuroscience that are both immature and lack stable and potentially valid explanatory targets in their own right is an extremely difficult task. Before considering the structure of one inter-disciplinary initiative that has these aims in mind, and in order to be able to evaluate its potential for achieving these aims, I want to consider some of the differences between the two areas of neuroscience that may be regarded as obstacles to achieving these aims.

Cognitive neuroscientists generally assume that "mental function is composed of distinguishable fundamental processes and that these processes can be selectively engaged by properly designed experimental task manipulations" and localized, via functional imaging techniques (e.g., fMRI, EEG), to specific areas of the brain (Carter et al. 2009, p. 169). In designing such task manipulations, cognitive neuroscientists engage in task analysis; they aim to provide "a clear specification of the [cognitive] processes thought to be engaged by the experimental task" and to determine "how these processes will be influenced by the variables to be manipulated in the experiment" (ibid.). For example, if an investigator is interested in creating a task to discriminate recognition memory from familiarity memory, in the process of designing the task she will consult previously published studies in the research literature, textbook understandings of each cognitive function, what is known about the temporal ordering of these processes in relationship to other processes, and which areas are thought to subserve those other processes (e.g., attention, working memory). Designing tasks that successfully individuate a cognitive function and simultaneously localizing that function in the brain requires a great deal of ingenuity. Cognitive neuroscientists have substantial freedom to produce, detect, and measure cognitive functions using the experimental paradigm or task that they take to be most reliable for achieving their investigative aims.

Not all cognitive neuroscientists will agree that a particular experimental task or paradigm is subject to only one unique task analysis. In fact, disagreements about the potential functions that play a role in the execution of a given task may prompt revisions to that task and/or the development of new tasks. Such disagreements have arisen in the cognitive neuroscience literature with respect to the Stroop task, which has been widely used to understand cognitive deficits in schizophrenia. (See, e.g.,

Perlstein et al. 1998; Barch et al. 1999a,b, 2004; Cohen et al. 1999.) In the standard version of the Stroop task, subjects are presented across trials variously with the word "red," the word "green," and a color-neutral word (such as "dog") in red or green type. The trials typically include congruent, incongruent, and neutral stimulus presentations. In the congruent condition the color of the text of the word matches the color word (e.g., "red" is presented in red-faced type); in the incongruent condition the color of the text differs from the color word (e.g., "red" is presented in green-faced type); in the neutral condition a color-neutral word is presented in either red or green type. The subject's reaction time from the point of presentation of the stimulus on a given trial to the point of responding with the correct word for the color seen (but not read) is measured, and errors in identifying the correct color word are recorded. On the non-computerized card version of the task, "patients [with schizophrenia] exhibit overall slowing of reaction times estimated across an entire card or . . . complete fewer items across all conditions" (Perlstein et al. 1998, p. 414). On the computerized version of the task, "normal" subjects have been shown to exhibit increased reaction-time (RT) interference (measured as difference between reaction time on neutral and incongruent trials) on the incongruent condition but exhibit few errors. Schizophrenics do not exhibit increased reaction time interference but may make more errors on the incongruent trials than controls and may also show increased RT facilitation (measured as the difference in neutral-congruent reaction time). Many investigators have interpreted the data as indicating that schizophrenics have a deficit in selective attention, insofar as they cannot "attend to a single dimension of a stimulus while simultaneously ignoring other task-irrelevant dimensions" (Perlstein et al. 1998, p. 414). Other investigators suggest that schizophrenics have a deficit in inhibition because they are unable to ignore the word in the incongruent condition, or because they "have difficulty suppressing the intrusive effects of the words" (ibid., p. 414). Another explanation is that schizophrenics have a deficit in "context processing" insofar as they cannot actively use information about context or what they are required to do in a task in order "to mediate task appropriate behavior" (Cohen et al. 1999). Thus, precisely what cognitive function the Stroop task individuates remains a subject of debate. (See also Barch et al. 2004.)

Such disagreements stem from several sources. First, there are different versions of the Stroop task. The computerized and non-computerized versions should not be classified as identical tasks, because the subject-stimulus interaction and the task demands placed on the subject (e.g., requiring

a subject to press a button or provide a verbal response) differ. However, even if we concede that these two versions of the same task are really different tasks, we still encounter, with respect to each version of the task on its own, differences in the number and ordering of stimuli presented and in the duration of inter-stimulus and inter-trial intervals. Such differences again prompt us to wonder if these tasks may be classified as the same and whether they can be said to involve the same cognitive processes and in the same ways.

The proliferation of competing hypotheses as to what cognitive function is disrupted in schizophrenia on the basis of Stroop-task data also makes sense in view of the fact that that different investigators put forward different task analyses to understand why schizophrenics exhibit the errors they do. Since these task analyses are likely to vary as a function of the version of the task that is being analyzed, we can anticipate that there will never be one single analysis of the Stroop task.

The Stroop task is also a complex task insofar as it places demands on a variety of what historically have been regarded as separable cognitive processes, including attention, working memory, language processing, visual processing, and implicit memory. For this reason, there is increasing consensus among cognitive neuroscientists that the Stroop task is not effective for individuating a single cognitive function and that Stroop-task data may not be used to adjudicate between competing interpretations of the cognitive functions it involves and the precise functions that are disrupted in schizophrenia.

Which areas of the brain are involved in Stroop-task performance has also been a focus of debate, different versions of the task being accompanied by different activation profiles for prefrontal cortex and anterior cingulate. (See, e.g., Adams and David 2007.) The fact that the behavioral data and the imaging data are subject to multiple distinct interpretations suggests that the Stroop task, which historically has been one of the better tasks for studying cognitive dysfunction in schizophrenia, does not individuate a single stable cognitive function or allow localization of a function in the brain in a way that could direct investigations into the cellular and molecular mechanisms that mediate such functions.

Such differences may be attributable to differences in task design across different computerized version of the task, but they may also be attributable to differences across research studies in the criteria used to recruit schizophrenic subjects. For example, the diagnostic screening and interview instruments or rating scales (see, e.g., Rush, First, and Blacker 2008) used to recruit schizophrenic subjects are not standardized across

laboratories. Some investigators take satisfaction of the *DSM*-IV definition alone to be sufficient; others require subjects to have a specific score on the Schizophrenia Rating Scale. And whereas some investigators exclude subjects on the basis of comorbidity of schizophrenia with another mental disorder (e.g., depression), a history of substance abuse, or a low IQ, other investigators use some of these criteria or none of them.[1]

Overcoming obstacles such as the ones identified above is a primary aim of the CNTRICS initiative. In fact, progress in cognitive neuroscience more generally is thought to be attainable by refining experimental tasks and correspondingly the taxonomy of kinds of functional processes as well as by standardizing fundamental features of experimental tasks and protocols. (See, e.g., Carter et al. 2009.) If such measures are successful, functional localization claims could become ever more fine grained, increasing the possibility of discovering the cellular and molecular mechanisms of these functions. Thus, cognitive neuroscience may be regarded as a field aiming towards stabilizing its kinds.

However, stabilizing the kinds of cognitive processes under study in cognitive neuroscience brings neuroscience only part of the way toward identifying the mechanisms that give rise to cognitive functions. Though it tells us where in the brain to look for cellular and molecular mechanisms, it does not tell us what those mechanisms are. Inter-disciplinary initiatives such as CNTRICS will be successful only if analogous cognitive functions are identified and stabilized in animal models. That will require the development of experimental paradigms that can be used to produce, measure, and detect cognitive functions in animal models that are analogous to those cognitive functions disrupted in human schizophrenics. The advantage of animal models is that they allow investigators to intervene in cellular and molecular activities in order to identify the mechanisms of cognitive functions and to determine the sources of cognitive dysfunction. They also provide a context in which to test the efficacy of pro-cognitive agents in improving cognitive dysfunctions before testing them on humans with schizophrenia.

However, as I have argued previously (Sullivan 2009), the forms of learning and memory under study in cognitive neurobiology are unstable for at least two reasons. First, similar to cognitive neuroscience, methodological pluralism is widespread in cognitive neurobiology. The experimental paradigms and adjoining protocols used to investigate the cellular and molecular mechanisms of cognitive functions or processes vary from one laboratory to the next. For example, two experimental paradigms that are both used to detect a cognitive function such as social recognition

memory may differ with respect to the type of stimuli used, the intensity or duration of the stimuli, and the duration of the interstimulus and intertrial intervals. Such differences could lead to differences in what cognitive functions are required for performance of a task and what cellular and molecular activities are involved. Thus, there are no real grounds upon which to establish that two labs that use different variants of an experimental paradigm are investigating the "same" phenomenon or its mechanisms, and this precludes integration of explanatory claims emanating from the two laboratories into a common explanatory model of the same function.

One reason we encounter what I have referred to as "a multiplicity of experimental protocols" in cognitive neurobiology is that different investigators have different intuitions about which constraints on the experimental process are most important. Some investigators are concerned with ensuring the reliability of the experimental process by using stimulus parameters that they are confident will produce data that will enable them to discriminate between competing claims about the effect under study in the laboratory. Often, however, such investigators sacrifice the external validity of their interpretive claims, because the stimulus parameters they select are not sufficiently similar to "real-world" stimulus parameters. Generally speaking, such methodological pluralism is encouraged in the hope that it may allow for novel findings that could not be achieved if experimental paradigms and protocols were standardized across investigators. However, such freedom is an impediment to the development of multilevel mechanistic explanations of mental disorders.

A second obstacle to stabilizing cognitive functions in cognitive neurobiology has to do with the fact that cognitive neurobiologists do not engage in task analysis when they design experimental paradigms to probe for cognitive functions of interest. As I demonstrated with respect to the cognitive function of spatial memory (Sullivan 2010), investigators are less interested in the cognitive processes that occur when an animal is trained in an experimental learning paradigm than with obtaining data indicating that an observable change in behavior has occurred—data that can be used as a basis for inferring that the cognitive function that the paradigm purportedly individuates has been detected. The trouble with this, however, as I demonstrated with respect to the Morris water maze, is that terms designating cognitive functions are often applied to sets of behavioral effects under study in an experimental paradigm in instances when no investigator is precisely certain what function the paradigm can be used to individuate. This leaves investigators somewhat free to liberally apply

different terms to refer to the function under study, to the extent that the term designating the function begins to oscillate. Such oscillations reveal that, at least in cognitive neurobiology, little work is done to understand what model organisms trained in experimental paradigms actually learn. The vast majority of cognitive neurobiologists are interested exclusively in the relationship between molecular changes and observable changes in behavior that they take to be indicative of a cognitive function of interest (e.g., changes in what has been learned or in memory). From the perspective of their immediate research interests, they don't care about the mental lives of their animal subjects. Thus, they don't worry that the experimental paradigms they use to probe for cognitive functions circumscribe many different functions, nor do they spend time worrying about how to modify experimental paradigms so that they track discrete functions. This is in stark contrast to cognitive neuroscience, which regards task analysis as a fundamental component in the interpretation of behavioral and imaging data and in the improvement of experimental paradigms.

I think it is safe to say, however, that neither cognitive neuroscientists nor cognitive neurobiologists are specifically concerned with the mental states of the organisms they study. Although cognitive neuroscientists do engage in task analysis, as far as I know none of them incorporate the potential mental or emotional states of their subjects into their explanatory models.

Given the kinds of differences that exist between cognitive neuroscience and cellular and molecular neurobiology, we can anticipate that the project of stabilizing mental disorders or cognitive functions and integrating results into multi-level explanatory models that reveal suitable targets for therapeutic intervention will fail. Those investigators who have decided to use the investigative tools on offer in cognitive neuroscience and cognitive neurobiology to identify the mechanisms that underlie mental disorders or cognitive dysfunctions will not succeed so long as there is no coordination across laboratories situated at the same and different levels of analysis to "stabilize the phenomena." I turn now to an evaluation of a research initiative that aims to create such coordination in order to assess its prospects and identify potential problems.

The CNTRICS Initiative

In 2007, two cognitive neuroscientists, Deanna Barch and Cameron Carter, described the development of an inter-disciplinary initiative whose ultimate aim was to develop psychopharmacological or "procognitive" agents

to "enhance cognition and functional outcome in schizophrenia" (Carter and Barch 2007, 1131). The Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS) initiative was in part a result of challenges faced by investigators involved in the Measurement of Treatment Effects on Cognition in Schizophrenia (MATRICS) initiative, a separate initiative spearheaded by Steven Marder and Michael Green in 2004. (See Carter and Barch 2007; Carter et al. 2008; Marder and Fenton 2004.) Investigators and pharmaceutical representatives who had participated in the MATRICS initiative, and the U.S. Food and Drug Administration, had the task of developing a battery of cognitive tests for the purposes of testing pharmacological agents to determine whether they improved cognition in schizophrenia. However, the battery of cognitive tasks that they compiled consisted primarily of "pen-and-paper-based" "clinical neuropsychological tests" that had been "developed for and validated in the clinical trials of atypical antipsychotics in the 1990s" (Carter and Barch 2007, p. 1132). Tasks then used in cognitive neuroscience to study cognitive dysfunction in schizophrenia were excluded, primarily because they were not thought to satisfy the criterion of construct validity. The lack of success achieved by MATRICS, primarily because the tasks selected failed to promote the discovery of pro-cognitive treatments, prompted a subset of the investigators familiar with or involved in the initiative to search for cognitive neuroscientific tasks that met the criterion of construct validity.

Cognitive tasks that had been widely used to study cognition and cognitive deficits in schizophrenia, such as the Stroop task and the Wisconsin Card Sorting Task, were acknowledged as involving multiple cognitive processes and were thus deemed "less helpful for understanding the specific nature of cognitive deficits, for identifying useful drug targets, [and] for assessing change in specific cognitive functions" (Carter and Barch 2007, 1133). As Carter and Barch note (ibid.), one ongoing aim of CNTRICS is to develop "process pure" tasks better capable of tracking single cognitive processes or subcomponents of more general processes.

In one of the initial research paper specifying the aims of CNTRICS, Barch et al. (2008) identified three other desiderata for candidate cognitive tasks—criteria that also correspond directly to the issues of "individuating" and "stabilizing" cognitive processes as scientific kinds. The first was consensus among cognitive neuroscientists as to which cognitive functions to investigate. They identified five broad categories of processes: "(1) executive control, (2) working memory, (3) long-term learning and memory (including reinforcement learning), (4) attention, (5) perception, and

(6) social and emotional processing" (Barch et al. 2008, p. 614). A working group was assigned to each cognitive function category, and each working group identified a set of sub-processes associated with that category for which experimental tasks that met the criteria already existed or could readily be developed. One subcomponent of attention identified for further study was control of attention, meaning "the ability to guide/change the focus of attention in response to internal representations" (Nuechterlein et al. 2009). One task that was designed to probe for this cognitive process was the "Guided Search Paradigm" (ibid.), a task in which a subject searches for a target in a visual array that differs from other stimuli in the array with respect to one feature (e.g., it is a square rather than a circle, or it is red rather than green). The subject is then presented with sets of visual arrays that contain the target and some distractors. The subject is asked to press a button to indicate whether the target is present. The advantage of such a task over the Stroop task is that it doesn't involve language processing, and hence it is likely that fewer cognitive processes are involved. Therefore, the task comes closer to potentially satisfying the "process pure" requirement and potentially allows for more precise localization of the cognitive function in the brain. Furthermore, the absence of a language component allows for easier translation of the task to animal models (Lustig et al., in press). Taken in combination, these features of the Guided Search Paradigm satisfy some of the conditions that I said had to be met for inter-disciplinary integration to be successful. First, it is important that a cognitive task individuate a discrete cognitive function; the simpler the task, the more likely it will be to serve this function. Second, investigators must agree that the task may be used to measure the function in question so that use of that task for measuring that function—the operationalization of that function—is standardized across investigators.[2] So, generally speaking, paradigms like Guided Search hold some promise for stabilizing the cognitive function of "control of attention." A second requirement, put forward by Carter and Barch (2007), is that the protocols associated with each cognitive task must be standardized across investigators. Carter and Barch acknowledge that this is primarily because different investigators "may use similar but nonidentical tasks to measure the same cognitive construct" and "these tasks . . . could . . . vary widely in potentially important characteristics such as number of trials . . . frequency of different trial types and the types of conditions included" (ibid., p. 1134). Imposing standardization of a task across investigators might eliminate problems associated with the multiplicity of experimental paradigms or protocols to test the same cognitive function at a single level of analysis. Furthermore,

such measures might ensure the stability of the referents of the terms designating specific cognitive processes, at least among cognitive neuroscientists.

The first three CNTRICS meetings yielded a battery of cognitive tasks, but the goal of the fourth through sixth meetings (2011–2012) was to identify imaging biomarkers—"characteristic(s) that are measured objectively as an index of a pathogenic process or a response to treatment" (Carter et al. 2011, p. 7)—that might be used to differentiate the normal brain from the schizophrenic brain by means of fMRI scans or EEG recordings undertaken during performance of the selected cognitive tasks. Investigators would then be able to determine the effects of pro-cognitive agents on the imaging biomarkers and on task performance (See, e.g., Carter et al. 2012a). The ultimate aim of identifying imaging biomarkers was to develop "an optical mechanism for translational research" so that "across levels of analysis . . . a common conceptual framework, language, and set of experimental tools that allows basic science to inform clinical and therapeutic research" could be applied (ibid.). However, so far little progress has been made in identifying imaging biomarkers for the cognitive functions that are of interest to CNTRICS investigators. For example, "basic and preclinical research has not progressed to the point where biomarkers of attentional control are fully ready for use in treatment research" (Luck et al. 2012, p. 59).

Though at least some of the aforementioned measures, taken together, may satisfy the requirements of stabilizing cognitive functions within cognitive neuroscience, they are not sufficient for stabilizing these functions within cognitive neurobiology and across the two areas of neuroscience. However, other criteria put forward by CNTRICS are intended to lay the groundwork for meeting that requirement. In addition to ensuring that the cognitive processes are impaired in schizophrenia and can be individuated in practice via appropriately designed experimental tasks, the CNTRICS working groups sought out tasks that can be used to "establish[] links with known neural circuits and neurotransmitter systems" and that have analogs that can be used in conjunction with animal models (Barch et al. 2008, p. 614).

The aim of the most recent meetings of CNTRICS has been to develop animal models for the purposes of screening drug treatments for the cognitive processes thought to be disrupted in schizophrenia. (See, e.g., Dudchenko et al. in press.) One task is to develop animal models of schizophrenia—in other words, rodent subjects that exhibit behaviors considered analogous to human subjects with schizophrenia. The second aim

is to identify tasks that are analogous to the tasks used to study cognitive deficits in schizophrenics. For example, investigators seek to identify a set of experimental paradigms that can be used to delineate "control of attention" in rodent models. Three tasks have been selected to probe for this cognitive function in animal models: the five-choice serial reaction time test, the five-choice continuous performance test, and the distractor-condition sustained-attention task. This third task most closely resembles the task conditions in the Guided Search paradigm (described above) that are designed to test control of attention in human subjects. Lustig et al. (2012, p. 1) provided the following reasons for selecting these three tasks:

> [T]he highest priority was given to construct validity, both in terms of the ability of the paradigm to specifically measure the process of interest and evidence that it recruited the neural systems thought to be critical for that process and impaired in schizophrenia. Reliability and the ability to standardize the paradigm across laboratories were also major concerns.

These criteria are exemplary of the kinds of strategies that I suggested could stabilize cognitive functions within cognitive neurobiology. If it can be established that all these paradigms measure the same cognitive function, then they also satisfy Wimsatt's (2007) criterion of robustness or the ability to access the same phenomenon via multiple different experimental procedures. Furthermore, the similarity between the Guided Search Paradigm to be used with human subjects and the distractor-condition sustained-attention task to be used with animal models increases the likelihood of being able to directly relate results across the two areas of neuroscience.

At best, insofar as no perfect animal model of schizophrenia exists, one might imagine the following kind of ideal scenario for stabilizing the phenomenon of "control of attention": Research studies using the Guided Search paradigm are conducted with an experimental group consisting of schizophrenic subjects and a control group consisting of "normal" human subjects. The two groups are compared on the performance of certain tasks, and the data are taken to indicate a deficit in "control of attention" in schizophrenic subjects. The study is replicated in different laboratories by different investigators. The results of those studies match the features of the original study in all relevant respects: the experimental paradigms, the experimental protocols, and the criteria used for recruiting schizophrenic subjects are the same. Meanwhile, a similar study is conducted using rodent subjects. An intervention technique (e.g., a lesion or a pharmacological manipulation) is used to produce a deficit in control of attention in one group of rats. The performance of that experimental group is compared against the performance of a group of "normal" rats on the distractor-condition sustained-

attention task. The data are taken to indicate that the rats in the experimental group exhibit a deficit in "control of attention." The experiment is replicated in several laboratories, with all the features standardized. A pharmacological agent is then introduced into the rats exhibiting a deficit in "control of attention." These experiments are also replicated across laboratories. When enough data on the pharmacological agent's efficacy in animal models have been accumulated (that is, when the drug meets FDA requirements), clinical trials using human subjects will begin.

This is my understanding of how investigators think the story will go for each of the functions identified as disrupted in schizophrenia, with candidate pro-cognitive agents being tested for each domain of cognitive function thought to be disrupted.

Prospects for Success of the CNTRICS Initiative

As I noted at the outset of this chapter, there is widespread agreement among philosophers of psychiatry that current categories of mental disorders fail to track stable regularities in nature and thus do not constitute natural kinds. There is also a growing consensus among research scientists that the *DSM*'s categories of mental disorders are not sufficient for grounding the search for causes of mental disorders and that a different classification scheme informed by measurement techniques that produce valid constructs is required. (See, e.g., Insel 2013.) The CNTRICS initiative and related NIMH-sponsored MATRICS initiative are the first such initiatives directed at eventually replacing categories such as schizophrenia and depression with a taxonomy designating cognitive functions as (for example) "control of attention" and "reward-based learning." This move prompts two questions: What will happen to the stability of mental disorders in this new system? How stable will the new scientific kinds that will replace them be?

The aim of the CNTRICS and MATRICS initiatives is not to stabilize current categories of mental disorders. If a new competing classification system of cognitive functions consisting of valid constructs begins to emerge and individuals begin to be diagnosed as having specific cognitive dysfunctions in addition to having mental disorders, we might anticipate that current categories of mental disorders could become wildly unstable and that such instability could compromise the stability of the new classification scheme. In other words, the new classification scheme, even with all its emphasis on the stability of kinds of cognitive functions in the form of valid constructs, may become subject to Hacking's looping effects.

This brings us to the question of whether the kinds of strategies of stabilization put forward by investigators involved in the CNTRICS initiative will ultimately be successful in stabilizing cognitive functions as explanatory targets—which may be considered an advantage over the *DSM*. One problem that we may foresee is, as I have demonstrated, that the areas of neuroscience that are trying to accelerate the stabilization of their explanatory targets (i.e., cognitive functions) are not mature sciences in Hacking's sense. The projects of localizing cognitive functions in the brain and identifying their cellular and molecular mechanisms are still in their infancy. To try to identify a set of cognitive tasks for each broad domain of function and its sub-functions and standardize them across research contexts immediately is likely to impede future scientific progress and prevent positive refinements to current taxonomies of cognitive functions. As philosophers of neuroscience (e.g., Bechtel and Richardson 1993; Bechtel 2008) have correctly pointed out, in the search for the mechanisms of cognitive functions the phenomena are likely to change or be "reconstituted" in light of new discoveries, and it is important that investigators remain open to this possibility. In other words, while stabilizing explanatory targets requires collective multi-disciplinary efforts such stabilization is likely only to emerge very gradually over time, if at all. Furthermore, CNTRICS is not the only interdisciplinary initiative directed at understanding the causes of mental disorders. The moral of the story is that coordination across different laboratories and different levels of analysis is desirable for discovering the causes of mental disorders, but a pluralism that promotes different investigative strategies is preferable. The precise form that such pluralism ought to take will have to be saved for another occasion.[3]

Notes

1. These claims are based on a partial analysis of the experimental literature on the Stroop task that considered nine papers published in eleven years (Thoma et al. 2007; Levy et al. 2004; Kerns et al. 2005; Yücel et al. 2002; Alain et al. 2002; Barch et al. 1999a,b; Perlstein et al. 1998; Carter et al. 1997). I regard this number of research studies as sufficient for establishing differences in subject recruitment across research studies and investigators. A broader analysis of the experimental literature would simply put my claim on stronger footing.
2. The same basic strategy was involved in the development of tasks for each of the five categories of cognitive processes thought to be disrupted in schizophrenia (e.g., long-term memory (Ragland et al. 2009), working memory (Barch et al. 2009a), and executive function (Gilmour et al. in press; Barch et al. 2009b; Carter et al. 2012b),

with a variety of tasks being nominated for research in clinical trials. (See, e.g.,, Barch et al. 2009c, p. 111, table 1.)
3. At least some research scientists appear to be in agreement, insofar as it is indicated in the draft of the NIMH Research Domain Criteria Project, that the constructs that will result from initiatives such as CNTRICS and MATRICS will be "subject to continual refinement with advances in science" (NIMH 2011).

References

Adams, R., and A. David. 2007. Patterns of anterior cingulate activation in schizophrenia: A selective review. *Neuropsychiatric Disease and Treatment* 3 (1): 87–101.

Alain, C., H. McNeely, Y. He, B. Christensen, and R. West. 2002. Neurophysiological evidence of error-monitoring deficits in patients with schizophrenia. *Cerebral Cortex* 12 (8): 840–846.

American Psychiatric Association. 1994. *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition.

American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, revised (*DSM*-IV-TR).

Barch, D., C. Carter, P. Hachten, M. Usher, and J. Cohen. 1999a. The "benefits" of distractability: mechanisms underlying increased Stroop effects in schizophrenia. *Schizophrenia Bulletin* 25 (4): 749–762.

Barch, D., C. Carter, W. Perlstein, J. Baird, J. Cohen, and N. Schooler. 1999b. Increased Stroop facilitation effects in schizophrenia are not due to increased automatic spreading activation. *Schizophrenia Research* 39 (1): 51–64.

Barch, D., C. Carter, and J. Cohen. 2004. Factors influencing Stroop performance in schizophrenia. *Neuropsychology* 18 (3): 477–484.

Barch, D., C. Carter, A. Arnsten, R. Buchanan, J. Cohen, M. Geyer, et al. 2008. Measurement issues in the use of cognitive neuroscience tasks in drug development for impaired cognition in schizophrenia: A report of the second consensus building conference of the CNTRICS initiative. *Schizophrenia Bulletin* 34: 613–618.

Barch, D., M. Berman, R. Engle, J. Jones, J. Jonides, A. Macdonald 3rd, et al. 2009a. CNTRICS final task selection: working memory. *Schizophrenia Bulletin* 35 (1): 136–152.

Barch, D., T. Braver, C. Carter, R. Poldrack, and T. Robins. 2009b. CNTRICS final task selection: executive control. *Schizophrenia Bulletin* 35 (1): 115–135.

Barch, D., C. Carter, A. Arnsten, R. Buchanan, J. Cohen, M. Geyer, et al. 2009c. Selecting paradigms from cognitive neuroscience for translation into use in clinical

trials: Proceedings of the third CNTRICS meeting. *Schizophrenia Bulletin* 35: 109–114.

Barch, D., H. Moore, D. Nee, D. Manoach, and S. Luck. 2012. CNTRICS imaging biomarkers selection: Working memory. *Schizophrenia Bulletin* 38 (1): 43–52.

Bechtel, W. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. Erlbaum.

Bechtel, W., and R. Richardson. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. MIT Press.

Belger, A., and D. Barch. 2009. Cognitive neuroscience and neuroimaging in schizophrenia. In *Neurobiology of Mental Illness*, third edition, ed. D. Charney and E. Nestler. Oxford University Press.

Bogen, J. 1988. Comments. *Noûs* 22: 65–66.

Carter, C., et al. 1997. Anterior cingulate gyrus dysfunction and selective attention deficits in schizophrenia: [150]H2O PET study during single-trial Stroop task performance. *American Journal of Psychiatry* 154 (12): 1670–1675.

Carter, C., and D. Barch. 2007. Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: The CNTRICS initiative. *Schizophrenia Bulletin* 33 (5): 1131–1137.

Carter, C., D. Barch, R. Buchanan, E. Bullmore, J. Krystal, J. Cohen, et al. 2008. Identifying cognitive mechanisms targeted for treatment development in schizophrenia: An overview of the first meeting of the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia Initiative. *Biological Psychiatry* 64: 4–10.

Carter, C., J. Kerns, and J. Cohen. 2009. Cognitive neuroscience: bridging thinking and feeling to the brain, and its implications for psychiatry. In *Neurobiology of Mental Illness*, third edition, ed. D. Charney and E. Nestler. Oxford University Press.

Carter, C., D. Barch, and E. Bullmore. 2011. Cognitive neuroscience treatment research to improve cognition in schizophrenia and related disorders. *Biological Psychiatry* 70: 7–12.

Carter, C., D. Barch, and the CNTRICS Executive Committee. 2012a. Imaging biomarkers for treatment development for impaired cognition: Report of the sixth CNTRICS meeting: Biomarkers recommended for further development. *Schizophrenia Bulletin* 38 (1): 26–33.

Carter, C., M. Minzenberg, R. West, and A. Macdonald 3rd. 2012b. CNTRICS imaging biomarker selections: Executive control paradigms. *Schizophrenia Bulletin* 38 (1): 34–42.

Cohen, J., D. Barch, C. Carter, and D. Servan-Schreiber. 1999. Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology* 108 (1): 120–133.

Cooper, R. 2004. Why Hacking is wrong about human kinds. *British Journal for the Philosophy of Science* 55: 73–85.

Craver, C. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford University Press.

Craver, C. 2009. Mechanisms and natural kinds. *Philosophical Psychology* 22: 575–594.

Cronbach, L., and P. Meehl. 1955. Construct validity in psychological tests. *Psychological Bulletin* 52: 281–302.

Dudchenko, P., J. Talpos, J. Young, and M. Baxter. In press. Animal models of working memory: A review of tasks that might be used in screening drug treatments for the memory impairments found in schizophrenia. *Neuroscience and Biobehavioral Reviews*.

Gilmour, G., A. Arguello, A. Bari, B. Brown, C. Carter, S. Floresco, et al. In press. Measuring the construct of executive control in schizophrenia.

Green, M., J. Lee, and R. Kern. 2009. Neurocognitive Assessment for Psychiatric Disorders. In *Neurobiology of Mental Illness*, third edition, ed. D. Charney and E. Nestler. Oxford University Press.

Hacking, I. 1983. *Representing and Intervening*. Cambridge University Press.

Hacking, I. 1988. On the stability of the laboratory sciences. *Journal of Philosophy* 85 (10): 507–514.

Hacking, I. 1989. The life of instruments. *Studies in History and Philosophy of Science* 20: 265–270.

Hacking, I. 1992. The self-vindication of the laboratory sciences. In *Science as Practice and Culture*, ed. A. Pickering. University of Chicago Press.

Hacking, I. 1995a. *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Princeton University Press.

Hacking, I. 1995b. The looping effects of human kinds. In *Causal Cognition: A Multidisciplinary Debate*, ed. D. Sperber, D. Premack, and A. Premack. Clarendon.

Hacking, I. 1998. *Mad Travelers: Reflections on the Reality of Transient Mental Illness*. University of Virginia.

Hacking, I. 2007. Kinds of people: moving targets. *Proceedings of the British Academy* 151: 285–318.

Haslam, N. 2000. Psychiatric categories as natural kinds: Essentialist thinking about mental disorders. *Social Research* 67: 1031–1058.

Haslam, N. 2002. Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry, & Psychology* 9: 203–217.

Hor, K., and M. Taylor. 2010. Suicide and schizophrenia: A systematic review of rates and risk factors. *Journal of Psychopharmacology (Oxford, England)* 24 (4): 81–90.

Insel, T. 2013. Transforming diagnosis. At http://www.nimh.nih.gov/about/director/index.shtml.

Kendler, K. S., P. Zachar, and C. Craver. 2011. What kinds of things are psychiatric disorders? *Psychological Medicine* 22: 1–8.

Kerns, J., J. Cohen, A. Macdonald 3rd, M. Johnson, V. Stenger, H. Aizenstein, and C. Carter. 2005. Decreased conflict and error-related activity in the anterior cingulate cortex in schizophrenia. *American Journal of Psychiatry* 162 (10): 1833–1839.

Khalidi, M. A. 2010. Interactive kinds. *British Journal for the Philosophy of Science* 61: 335–360.

Kincaid, H. 2008. Do we need theory to study disease? Lessons from cancer research and their implications for mental illness. *Perspectives in Biology and Medicine* 51: 367–378.

Levy, D., N. Mendell, and P. Holzman. 2004. The antisaccade task and neuropsychological tests of prefrontal cortical integrity in schizophrenia: Empirical findings and interpretative consideration. *World Psychiatry; Official Journal of the World Psychiatric Association (WPA)* 3 (1): 32–40.

Luck, S., J. Ford, M. Sarter, and C. Lustig. 2012. CNTRICS final biomarker selection: Control of attention. *Schizophrenia Bulletin* 38 (1): 53–61.

Lustig, C., R. Kozak, M. Sarter, J. Young, and T. Robbins. In press. CNTRICS final animal model task selection: control of attention. *Neuroscience and Biobehavioral Reviews*.

Marder, S., and W. Fenton. 2004. Measurement and treatment research to improve cognition in schizophrenia: NIMH MATRICS initiative to support the development of agents for improving cognition in schizophrenia. *Schizophrenia Research* 72: 5–9.

NIMH. 2011. NIMH Research Domain Criteria (draft). Available at http://www.nimh.nih.gov.

Nuechterlein, K., S. Luck, C. Lustig, and M. Sarter. 2009. CNTRICS final task selection: Control of attention. *Schizophrenia Bulletin* 35 (1): 182–196.

Nuechterlein, K., K. Subotnik, J. Ventura, M. Green, D. Gretchen-Doorly, and R. Asarnow. 2012. The puzzle of schizophrenia: Tracking the core role of cognitive deficits. *Development and Psychopathology* 24 (2): 529–536.

Perlstein, W., C. Carter, D. Barch, and J. Baird. 1998. The Stroop task and attention deficits in schizophrenia: A critical evaluation of card and single-trial Stroop methodologies. *Neuropsychology* 12 (3): 414–425.

Ragland, J., N. Cohen, R. Cools, M. Frank, D. Hannula, and C. Ranganath. 2012. CNTRICS imaging biomarkers final task selection: Long-term memory and reinforcement learning. *Schizophrenia Bulletin* 38 (1): 62–72.

Rush, J., M. First, and D. Blacker. 2008. *Handbook of Psychiatric Measures*, second edition. American Psychiatric Publishing.

Sullivan, J. 2009. The multiplicity of experimental protocols: A challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese* 167: 511–539.

Sullivan, J. 2010. Reconsidering "spatial memory" and the Morris water maze. *Synthese* 177: 261–283.

Sullivan, J., and E. Thiels. 2011. The place of the mind in contemporary neuroscience. Presented at annual meeting of Society for Neuroscience, Washington.

Taylor, S., and A. MacDonald 3rd. 2011. Brain mapping biomarkers of socio-emotional processing in schizophrenia. *Schizophrenia Bulletin* 38 (1): 73–80.

Thoma, P., D. Zoppelt, B. Wibel, and I. Daum. 2007. Context processing and negative symptoms in schizophrenia. *Journal of Clinical and Experimental Neuropsychology* 29 (4): 428–435.

Uttal, W. 2003. *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. MIT Press.

Wimsatt, W. 2007. *Reengineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Harvard University Press.

Yücel, M., C. Pantelis, G. Stuart, S. Wood, P. Maruff, D. Velakoulis, et al. 2002. Anterior cingulate activation during Stroop task performance: A PET to MRI coregistration study of individual patients with schizophrenia. *American Journal of Psychiatry* 159 (2): 251–254.

Zachar, P. 2000. Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, & Psychology* 7: 167–182.