“That Proves You Mad, Because You Know It Not”:
Impaired Insight for Psychosis and the Liberal Dilemma of Madness

Candido: What? Am I mad, say you, and I not know it?
Officer: That proves you mad, because you know it not.
-  The Honest Whore, Thomas Dekker 1604

A majority of individuals with schizophrenia have poor insight regarding the fact that they have a psychotic illness. Evidence suggests that poor insight is a manifestation of the illness rather than a coping strategy. It may be comparable to the lack of awareness of neurological deficits seen in stroke, termed anosognosia.
-  Diagnostic and Statistical Manual IV TR, American Psychiatric Association 1994

Introduction:

For classical liberal theorists the mad served as a limiting case in the project of expanding rights. As John Locke ([1690] 1980 p.350) put it, “Madmen, which for the present cannot possibly have the use of right Reason to guide themselves, have for their Guide, the Reason that guideth other Men which are Tutors over them...” J.S. Mill ([1869] 1999 p.10) similarly argued that liberty was “to apply only to human beings in the maturity of their faculties,” and that “good despots” remained necessary for those who could not be educated into responsible freedom. Other disenfranchised groups might be trained into self-governance, but what of those whose very nature seemed to preclude such transformation? The paternalistic asylum of the 19th century offered a practical if not elegant solution to this dilemma, with caretakers deploying strategies from brute restraint to moral treatment in an effort to manage and normalize the mad (see Foucault 1965, Porter 2002).

The mid-twentieth-century closing of the asylums fundamentally altered the matrix of responsibility for and authority over mad people. The logic of democratization, so prevalent in the social movements and organizational changes of the era, proved an awkward fit in the case of psychiatric patients. Far from resolving the liberal dilemma of madness, deinstitutionalization and the civil libertarian turn in mental health law merely brought more questions to the fore: When should a person in the throes of psychosis be held criminally responsible for violation of law? Could the mad be trusted to make treatment decisions? Finally, in the wake of social scientific critique of psychiatry, how could a society legitimately determine who is so mad as to be in need of Mill’s “good despots”? Rather than pursue these questions through the normative lens of ethics or political theory, this article draws on sociological approaches to knowledge, classification, and power to investigate how actors and institutions have attempted to deal with the liberal dilemma of madness as a practical problem.

Psychiatric researchers have offered one rhetorically powerful and controversial solution to these questions by locating patient self-knowledge, and with it volition, in the brain. “Impaired insight for psychosis,” analogized to the neurological condition “anosognosia,” describes a person unable to know that he or she is mad and therefore incapable of rationally coping with that madness. Unlike previous theories of denial, which posited a psychodynamic defense mechanism for avoiding knowledge of illness,
the neurological model throws even unconscious volition into question. The dilemma is now put as follows: can an untreated person make a real choice regarding voluntary treatment if the disease itself impairs awareness of compromised rationality?

Using a genealogical approach (Foucault 1977), I investigate the historical conditions of possibility for the emergence of impaired insight as a scientific object, its linkage to compliance behaviors, and its translation into the legal and policy realms. Although reference to the mad who “know it not” precedes even the modern notion of mental illness, the scientific and political interest in impaired insight is a historically contingent product of democratization in the liberal welfare state. Within new legal-institutional environments that privileged patients’ negative liberty and predicated capacity on volition, impaired insight offered a way of conceptualizing and addressing emerging practical dilemmas tied to patient choice. Scholars cannot understand the development of insight science without analysis of changes in liberal institutions of governance. Less obvious is the role the classification plays in current efforts to reform those very liberal institutions, by shifting conceptions of agency and capacity, bridging actors across social fields, and offering a powerful frame for legislative mobilization.

In order to illuminate these linked transformations, I proceed through my empirical materials with a novel framework I call classificatory niche construction. My approach builds on Ian Hacking’s (1998) ecological niche approach to mental illness categories, which eschews hardline realism and vulgar constructivism by outlining historical conditions of emergence and life. After elucidating the problematization of impaired insight for psychosis, I use the case to demonstrate how Hacking’s niche metaphor can be made both more dynamic and explanatorily powerful. I do this first by expanding the empirical scope of his analysis “down” to micro-practices of technical scientific perception and “up” to institutions, and second, by keeping the ecological metaphor but updating Hacking’s older model of ecology with contemporary approaches focused on niche construction.

Much as ecologists altered standard evolutionary theory by studying the way organisms reconstruct their niches and subsequently themselves (see Lewontin 1978, Odling-Smee et al. 1996), I extend Hacking with tools from science studies, showing how a changing classification facilitates the remaking of its material and institutional ecology through materialization and translation processes. Situating Hacking’s classificatory niches within the rich tradition of ecological thought in sociology, I offer a relational approach to psychiatric labeling that is simultaneously attuned to techno-scientific, interactional, and institutional dimensions.

The study is oriented around two empirical puzzles:

1. Why did denial of illness emerge as an object of sustained psychiatric inquiry in the 1970’s, given that awareness of madness had long been noted as a variable dimension of psychosis?
2. Why did the concept, reframed as “impaired insight,” become relevant to American psychiatric law and policy in the 2000’s, given longstanding interest in reform of commitment procedure vis-à-vis patient competence?

The logic of my answer proceeds as follows:
The mid-twentieth-century shift in psychiatric patient management from institutional confinement to largely inadequate community care created a host of new dilemmas, drawing attention to psychiatric outpatients’ legally empowered refusal or failure to self-govern in prescribed ways. In this “ecological niche” the concept of insight became an object of scientific interest as a mediator of compliance, but the apparent volition contained in the psychoanalytic theory of denial was ill suited for translation into a legal paradigm predicated on volition. With the paradigm shift from psychoanalysis to biological psychiatry and development of new representative technologies, however, insight became scientifically measurable and visually perceptible as material pathology within the brain. With somaticization in the 1990’s insight became a “boundary object” (Star and Griesemer 1989) linking psychiatry, neurology, law, and policy, enabling new possibilities for civil commitment reform and changes to the ecology. Champions from the increasingly powerful family and treatment advocacy social movements then brought insight science to bear upon the very questions of custodial management, patient rights, and governance that gave birth to it.

To echo Valverde’s (1998) writings on choice and determinism in theories of alcoholism, the case of impaired insight is not merely the story of a medical problem, but of a paradox at the center of liberal governance. This paradox is that human beings are born free, but must also be made free (see Hindess 1993). At stake is the question of how societies that simultaneously privilege individual responsibility and somatic accounts of behavior will govern those at the borders of legal capacity, and with what justification. Thus, I use the case of insight not only to extend existing sociological theories of power, knowledge, and classification, but also to illuminate the historical conditions of possibility for a debate in liberal theory that has for too long been taken for granted.

Political philosophers have looked to Locke and Mill’s writings on madness to theorize the limits of rights holding, but found such statements vague and often self-contradictory (see Failer 2002). The central premise of this article is that efforts to reframe, resolve, or simply paper over the liberal dilemma of madness will not be understood through an exegesis of normative political theory, but through an empirical sociological analysis of the actor-network of clinicians, researchers, patients, psychological scales, brain scans, experimental interventions, and case law that have made mad peoples’ agency and insight an object of scientific truth. I turn now to an overview of sociological theories of classification and power, and outline the approach I will use to explain the case at hand.

From Labeling Theory to Dynamic Nominalism

Since the classics, sociologists have been concerned with understanding the origins and consequences of classificatory systems, and specifically how these come to be naturalized. From Marx’s (1867) reification, to Durkheim and Mauss’ (1905) primitive classification, to Berger and Luckman’s objectivation (1966), such theorists have argued that designations are most powerful when human input is obscured. The Chicago School produced the earliest focus on the consequences of human labeling, such as Tannenbaum’s (1938) finding that “tagging” youth as delinquents alters both a person’s self concept and the way others view him or her, leading to self-fulfilling prophecies of criminality. Numerous thinkers expanded this approach to encompass all
sorts of behavior, offering a radical social constructionist approach to deviancy more generally (see Lemert 1951, Erikson 1962, Becker 1963). As fruitful as such an agenda was, the application of labeling theory to the case of mental illness soon proved highly controversial.

Major works from Goffman (1961) and Scheff (1966) on the labeling roots of psychosis, alongside Foucault’s (1964) historicization of taken-for-granted medical formulations, and Laing’s (1967) insistence on the sanity of madness in an insane world, offered a radical challenge to prevailing psychiatric theories. If madness was indeed only a label used to mark deviance, then psychiatric diagnosis did not identify natural kinds so much as naturalize power relationships. Libertarian psychiatrist Thomas Szasz (1961) produced the clearest account of the liberal political implications of labeling theory, arguing that unless psychiatry could identify pathophysiology, diagnosis served merely to medicalize sociopolitical or moral distinctions. Going beyond the classical liberals, Szasz argued that paternalistic treatment constituted an unjustified deprivation of freedom under the guise of science.

Although popular during the 1960’s and 1970’s, and serving as one of several justifications for the deinstitutionalization movement (see Scull 1977) and a legitimacy crisis in psychiatry (Kirk and Kutchins 1992), labeling theory was soon dismissed for failing to reckon with the reality of psychiatric disability. Psychiatrists (Spitzer 1975) and even sociologists (Gove 1975) accused such thinkers of vulgar idealism and naïveté, and with ensuing transformations in psychiatric science and the explosion of the pharmaceutical industry, social reductionism soon gave way to the biological reductionism of neo-Kraepelinian psychiatry (Luhrmann 2000).

Contemporary medical sociologists have largely abandoned the etiological claims of labeling theory, which assumed that the primary effect was a self-fulfilling-prophecy that could create madness. Proponents of modified labeling theory (Link et al. 1989, 2001), for instance, investigate the effects of labeling on experiences of stigma and threats to self-esteem rather than the creation of mental illness. Horowitz and Wakefield’s (2007) research on the putative epidemic of depression offers a moderate constructionism, arguing that existing psychiatric taxonomies fail to differentiate between the real disorder of major depression and normal sadness improperly coded as depression. Others (see for instance Eyal et al. 2010) have turned instead to philosopher Ian Hacking’s (1986) “dynamic nominalism” for a re-theorization of psychiatric phenomena as emergent and interactive.

Hacking’s approach is concerned with the historical conditions of possibility for “human kinds” to come into being and change over time. Rather than take psychological and psychiatric categorizations as real, historical constants or the mere medicalization of deviant behavior, this approach takes such phenomena as “ontologically emergent, simultaneously real and historical” (Sugarman 2009, p 7). Human categorizations are not static representations of the real, but instead subject to “looping effects” (Hacking 1995) whereby classifications affect the classified, who, through subsequent action and interaction with institutions, remake themselves and the classification. Beyond the assumed self-fulfilling-prophecy of classical labeling theory, Hacking surmises that the classification process can have myriad effects in interaction with both humans and institutions. In his inquiries into the origins of mental illness categories, Hacking (1998) offers the metaphor of “ecological niches,” or receptive environments in which new or
“transient” illnesses can come to life, explaining how a disorder might require cultural conditions in addition to biological ones.

The ecological approach is rooted in his larger project to temper those aspects of Foucauldian genealogy that realists find most offensive. If genealogical studies sometimes investigate the historical conditions of an entity’s emergence by asking when it becomes possible to think something, Hacking’s metaphor is an attempt to move beyond the realm of discourse and show the concrete conditions through which it becomes possible to do something with a psychiatric phenomena, such as perform it, measure it, diagnose it, or treat it. In Mad Travelers (1998 p.82), a study of the transient mental illness of fugue, he sought to distance himself from the “Anglophone Foucault industry” and its undue focus on textual representation. If Hacking is dismissive of studies that tend toward linguistic idealism,

It is because discourse does not do the work. Of course language has a great deal to do with the formation of an ecological niche, but so does what people do, how they live, the larger world of the material existence that they inhabit. (Hacking 1998 p. 86. Emphasis added.)

He outlines four vectors, or niche conditions, in which a new disease can emerge with its classification. First is medical taxonomy, which requires that a disorder be recognizable within an existing framework of classification. Second is observability, that a potential disorder be intelligible and visible to authorities as a social problem. Third is cultural polarity, requiring a disorder be morally situated somewhere between dominant notions of virtue and vice. Fourth is release, which describes how undesirable disorders may offer those so labeled some sort of relief not available in the wider culture.

The social scientific reception to Hacking’s work has been mixed. Although there has been fruitful use of the looping concept and ecological niche metaphor in studies of such disease classifications as autism (Eyal et al. 2010) and social anxiety disorder (Hickinbottom-Brawn 2013), critics like Kuorikoski and Poyhonen (2012) suggest that the philosopher should become a sociologist and focus on social mechanisms rather than ontology. They accuse him and others of taking looping effects as “an argument against a realistic understanding of the human sciences,” (195) and aim to resituate the phenomenon in analytical sociology. A mechanism-based approach, they argue, could replace “hopelessly vague theorizing about the social construction of the phenomenon with explanatory causal accounts of the feedback-mechanism…” (Kuorikoski and Poyhonen 201).

One goal of this paper is to show that Hacking’s approach can coexist with the more boldly causal language of mechanisms, to the benefit of each perspective, but this requires a considerable expansion of Hacking’s empirical scope. In order to add plausible causal accounts to dynamic nominalism, I begin by opening the empirical frame of Hacking’s theories “down” to technical scientific questions of perception and “up” to the scale of institutional analysis. I then move emphasis from the ecological opportunities afforded and constraints placed on entities to the ways those entities remake their institutional environments.

From Niche Conditions to Niche Construction
While Hacking offers rich historical detail in his account of a particular niche, he is maddeningly vague in theorizing what the overarching ecology consists of, referring to it as the “environment” or “culture.” The ecological orientation is generative because it helps us see beyond polarized realist and vulgar constructionist approaches to psychiatric classifications, but there are two primary problems that prevent a full realization of this approach: First, by conceptualizing ecologies in the global terms of “culture” and “environment,” Hacking misses out on powerful empirical tools for capturing how micro-level technical scientific practices and meso-level institutions shape classified human kinds. These empirical concerns in turn point the way to a far more specific theory of niche conditions. Second, he has the right metaphor but the wrong model of ecology. For the founder of “dynamic nominalism, his model is surprisingly not dynamic, assuming a unidirectional set of pressures from the ecology onto the entity. Rather than standard evolutionary theory, the phenomena calls for a theorization of niche construction, which can model the way a classification remakes various elements of its environment and subsequently itself.

Ecological metaphors have long served as intermediate accounts of social thought between methodological individualists, who envision atomism, and structuralists, whose metaphor is functionally integrated organism. Ecology suggests “interaction between multiple elements that are neither fully constrained nor fully independent” (Abbott 2005 p 248). To think of classified human kinds as within an ecology suggests that they are neither simplistic biological natural kinds nor constructions called into being by a functional cultural system, but realities that are nonetheless dependent upon interaction with an array of other social and cultural phenomena for survival. When I describe the “ecology” for human kinds, I mean a set of relationships that can be empirically traced by following documents, diagnostic instruments, or practices across institutions and social fields. Where many ecological metaphors assume a single type of entity within ecologies (i.e. organizations in a field, people in a city, etc.), my ecologies contain classified human kinds, material infrastructures, reified scientific instruments, symbols, and a host of other entities that act in relation to both humans and each other.

Abbott (2005) has argued that ecological models too often bracket important factors outside of an immediate ecological frame as either static or simply unaccounted for. While analysts can often point to causally important conditions or pressures, there is a tendency to specify the ecology around a singular type of actor such that other forces are merely ignored. From Chicago School ethnographers who have little to say about how political economic factors affect cities, to world systems theorists who assume states are locked in ecological relationships but ignore other macro-ecological aggregates like religions, social ecologists miss the opportunity to show much more complex relations between various types of entities or even between types of ecologies. Falling into the same trap but from the other direction, Hacking’s ecology as “culture” is so broad that it fails to tell us about the smaller elements within the niche, larger ecology outside of the niche, or how that ecology relates to its surrounds.

The first task for remaking Hacking’s ecological approach, then, is to extend the empirical scope of each of his niche conditions “down” to micro-practices of perception.
and “up” to locate the niches in institutional ecologies. Here is where existing empirical tools from science and technology studies that address the material conditions of possibility for phenomena to become perceptible and then actionable—what Michele Murphy (2006) has called the “regime of perceptibility—and institutional theory’s conceptualizations of logics, field effects, resource dependencies and other pressures shape possibilities in the niche, become crucial.

In this light, Hacking’s (1998) niche conditions of 1) medical “taxonomic” systems is the result of both techno-scientific practices within laboratories and clinics, as well as larger institutional factors that determine funding streams and set research agendas. 2) “Observability” is a function of both technical perceptibility and institutional factors like which populations are defined as deviant and brought under medical surveillance to begin with. 3) “Cultural polarity” similarly becomes far more empirically tractable when it is re-specified as conflicting institutional logics within particular organizational fields that lead to concrete dilemmas of practice, versus a totalizing account of moral ambiguity within a “culture.” 4) “Release,” of course, only makes sense within the social role demands of particular institutional settings, like familial obligations, markets, or the state.

The second task is to open Hacking’s model to dynamic change, for even with a more empirically hefty account of niche conditions, Hacking’s transient illnesses are treated mostly as an outcome, with little causal power of their own to change the environment. The most significant omission in Hacking’s ecological niche model is the failure to theorize how entities remake their immediate environments and the larger ecological system, and subsequently themselves. Within contemporary ecology, entities do not emerge de novo from novel niche conditions, as in Hacking’s (1998) story of “traveler’s fugue,” but develop from previous forms in a process of mutual fitting between organism and environment. The changed environment may affect selection pressures on the entity, whether organism or classified human kind, which may adapt its form and behavior, once again altering the ecology, and so on and so forth in a dynamic feedback loop.

Hacking’s model elides the complexity of ecological thought because it draws on standard evolutionary theory, which assumes the top-down primacy of environmental conditions in the development of species. Here entities either die or adapt in response to ecological change, with little input from the organisms themselves. Lewontin (1978) has asserted, in contrast, that environments are themselves partially the intended and unintended consequence of organisms’ actions, and that organism-driven ecosystem

---

1 I lack the space to compare Hacking to every other use of the metaphor in sociology, but to put it simply, they are not all up to the same thing. Hannan and Freeman’s (1977) work on population ecology is less useful here because that paradigm is focused primarily on organizational birth and death, rather than transformation. See Singh and Lumsden (1990.) Baum and Powell’s (1995) call for an “institutional ecology” that gets at the “co-evolving nature of cultural understands, organizational forms, and resource constraints” is closer to how I’m thinking of process. I am highly sympathetic to ecology in Chicago School sociology, from the investigations of occupations (Hughes 1971), to interactions (Goffman 1963), to cities (Park, Burgess and MacKenzie 1925). In general, these approaches have not directly theorized “niche construction” as such.
engineering may in turn affect selection pressures on those organisms. Odling-Smee (1988) later termed this process “niche construction,” and from leaf cutter ants cultivating fungi, to beavers building dams, to humans laying transport infrastructure, ecologists now take seriously the *mutual fitting* between the organism and ecology. This perspective leads to an incredibly dynamic model of ecology, which must take account of multiple organisms’ niche construction, environmental feedback on behaviors and practices, as well as phenomena exogenous to the immediate ecology, like meteor showers.

Drawing on empirical analytic tools from science studies, and contextualizing niches within an institutional order, I offer *classificatory niche construction* as the kind of materially accountable genealogy Hacking (1998) proposed but did not deliver. The niche construction of leaf cutter ants and psychiatric phenomena are clearly different, but the theoretical construct is helpful because it points attention to certain dynamics that otherwise remain unexamined. The “mutual fitting” I focus on here consists of two interrelated processes, *materialization* and *translation*, through which an emergent entity like a new symptom or illness classification can in turn affect its symbolic and material ecology, and subsequently itself.

By materialization I mean the way in which thought and practice become stabilized through standardization (Latour 1987, Timmermans and Berg 1997), as when a cluster of disparate experiences is made transportable across time and space as a diagnosis, or when an abstraction becomes an actual material object via technologies like brain scanning (Dumit 2004). By translation (Callon 1985, Latour 1987), I mean the way in which such materializations cross borders between social fields, with facts sometimes stabilized and sometimes reconstructed, enrolling disparate actors in shared projects. Materialization processes can be understood as the response of the entity to selection pressures and changes in the ecology, whereas translation is the entity’s niche construction and remaking of the environment. Humans drive these processes, of course, but much of the human agency is only possible within a broader network of concepts, standardizations, literature, technologies, and legal code (Latour 1987).

The benefit of my *niche construction* approach is that it takes us from unidirectional conditions of emergence to a dynamic multidirectional feedback system of mutual fitting between entity and ecology. Hacking’s focus on looping through the mind and actions of the classified person is an obviously important form of feedback, but it ignores the multitude of ways in which a new human classification can affect the symbolic and material foundations of its ecology. Classifications do not only feedback through classified individuals changing their behavior, but also through the creation of instruments, representations, and interventions, which can enable subsequent translations and institutional changes to the broader symbolic and material ecology, in turn triggering more materializations and translations of the entity. By bringing these empirical analytic tools into Hacking’s project, I can concretely investigate how ideas, objects, and entities interact with the larger social, material, political environment as an ongoing process, rather than as a singular moment of emergence.

**Methods and Data Collection:**

This article is part of a larger project examining the reconfiguration of American psychiatric science, community mental health and welfare state practice, and patient-
during the 20\textsuperscript{th} and early 21\textsuperscript{st} century. The argument put forth here draws on historical materials, including scientific journal articles, standardized medical scales and measures, law review articles, case law, chapters in edited volumes, blogs, the so-called “pink literature” of medical advocacy groups, websites of organizations, popular news media, and legislation. Using a genealogical approach (Foucault 1977), I investigate the historical conditions of possibility for the emergence of insight as a scientific object, its linkage to compliance behaviors, and its translation into the legal and policy realms.

Following methodological developments in the socio-historical study of science, I aimed to unravel concepts and expertise that have been “black boxed” (Latour 1987), or rendered invisible by their own success and smooth functioning in technical and practical projects. As Eyal (2013b p. 871) has described it, the construction of scientific facts and expertise may be seen early on, when “alternative devices, actors, and arrangements are still viable candidates for formulating the problem or addressing it.” After this period of visible contingency, however, scholars must begin with the reified concept and work backwards, so that the process of ‘black boxing’ and standardization can be reversed to trace the other actors (even if they do not control any identifiable part of the task at hand), equipment (especially transcription devices), institutional and spatial arrangements, and concepts of which a form of expertise is composed (2013b p. 872).

After first tracing the black boxing of, say, a psychological scale to the conditions under which the scale was constructed and used, I next followed where the black box travels and whether the meaning of the measurements changes with new uses.

While quantitative approaches to diffusion can powerfully show how a singular concept travels across fields, it is precisely the object’s transitions and stabilizations in relation to larger societal dynamics that are at stake here. Given that the phenomenon has alternately been formulated as “poor insight,” “illness awareness,” “impaired insight,” “the wish to be crazy,” “denial,” and “anosognosia,” a close reading of all materials was necessary to trace the connections between disparate terminology and to avoid conflating distinct approaches that share keywords. Psychoanalysts, for instance, frequently address “insight” as self-knowledge in general, and neurologists often discuss “anosognosia” in reference to stroke victims, so keyword search counts can be misleading. I therefore follow Latour (1987) in tracing translations rather than diffusion, for the assumption that a stable entity diffuses between pre-constituted social actors can obfuscate the way both the entity and the actors are themselves unstable and sometimes reconstituted in the process.

I began by searching PubMed for the terms “insight” and “psychiatry,” and then possible synonyms like those above and associated concepts (i.e. “non compliance”) that emerged as the analysis developed.\textsuperscript{2} I initially scrutinized texts for the context and the

\textsuperscript{2} A key tension in the social study of scientific literatures is that the corpus may actually be too large for an individual to read, yet change in such ways over time that textual analysis is preferable to citation analysis (see Kelty and Landecker 2009). A certain amount of filtering becomes necessary when a term like “insight,” as a lay and scientific word, has 122,940 entries on Pubmed alone. Even “insight” and “psychiatry” reveals 3292 entries, a daunting number for a close textual analysis. As I detail, my subsequent
meaning of those terms, and coded them by theoretical orientation and method (i.e., psychodynamic versus biological, case study versus statistical) and publication date (see Armstrong 2003 for a similar genealogical approach with scientific databases). By mapping the references by time, theoretical approach and valence, I could see, for instance, that scientists in the 1990’s continued to invoke the psychoanalytic denial theory, but in a negative rather than positive light. Other scientific boundary work (Gieryn 1983) helped me structure future searches with more specific terms and time frames. When I found scientists claiming a conceptual breakthrough in the turn to neurology in the 1990’s, for instance, I next searched neurological journals for the analogy between poor insight and “anosognosia” in earlier periods to see if such a thing had been previously thinkable. This revealed that in 1974 a neurologist had already considered the analogy but rejected it, another situation where citation or keyword counts might miss the meaning at hand.

In regards to relevant law review articles, case law, and briefs, I performed similar operations with the legal database Westlaw, organizing documents by the notion of insight used, the kind of legal argument in which it was deployed, and time period. Blogs, advocacy group statements, and other texts were found through google searches. In other cases I identified key authors and combed through publication records for expert statements occurring outside of the bounds of academic, professional or legal spheres.

As a further strategy, I utilized highly cited key documents and moved backwards in time by reading works referenced, and then forwards by reading newer articles citing the writings in question. In some cases I read through authors’ previous works to confirm suspected influences that were not acknowledged in the document at hand, or followed chains of citations revealing intellectual debts that may not have been known to individual authors themselves. Thus, my searching, coding, and analytic schemes were both inductively and deductively derived. Given the shifting meanings of key terms at different times and in different contexts, the inclusion of particular historical materials required interpretation from the outset, and the organizational coding was ultimately used to facilitate close reading. In the next section I begin by discussing early folk and scientific conceptualizations of insight in relation to madness.

Problematizing Insight: Outpatient Drug Refusal and the “Wish to be Crazy”

The notion of the putative mad person unaware of madness has been in circulation for some time, as evidenced by the character Candido in Dekker’s 1604 play “The Honest Whore.” For 17th century philosophers like Hobbes and Locke, a delusional person aware of delusions constituted a logical contradiction, with a lack of insight essentially definitional to madness (Berrios 1994). Late 19th and early 20th century theories allowed for a degree of patient self-knowledge, with Kraepelin (1917) viewing some dementia praecox patients as partially aware of illness before descending into further madness. Lewis (1934) defined insight as “correct attitude toward morbid change in oneself,” and seemingly used patients’ partial insight to affirm their basic humanity in a culture rife with eugenic pressures (according to David 1999 p. 215). Researchers expressed
empirical interest in insight in relation to the problem of “post psychotic depression,” observing that once patients gained full understanding of their madness they were forced to reckon with the dour reality of the situation and became sad (see Mayer-Gross 1920, as cited in McGlashan-Carpenter 1976).

Although a full institutional accounting of these early statements regarding insight is beyond the scope of this paper, it is useful for showing the contrast with the 20th century transformation. Consider that the mad people of Dekker’s (1604) era were treated as a domestic or local parish problem, for there were no asylums to speak of. Hobbes and Locke wrote in the earliest years of the English asylum, and took for granted that such people required paternalistic management. That lack of insight was definitional to madness was derived from first principles and speculation, for they lacked the raw empirical materials (locked up people) for observation. It was not until the “great confinement” (Foucault 1965) of the 18th century that the poor, unemployed, mad, and other deviants were brought into hospitals that medical doctors could begin to observe, classify, and theorize a science of madness.

When Kraepelin, Mayer-Gross and Lewis were writing in the late 19th and early 20th century, mad people throughout Europe and the United States were for the most part medically treated, segregated, and studied in asylums. To situate the relation of insight to “post-psychotic-depression,” then, consider that Mayer-Gross’s patients of the early 20th century would exit a psychosis and likely be confined in a total institution. It is perhaps unsurprising that clinicians and researchers saw individuals gain “insight” into their situation and find it depressing. Despite muckraking criticism of asylum conditions, however, the state’s parens patrie power to confine mad people for their own good went largely unquestioned in either the classical era or the early 20th century.

For the burgeoning discipline of psychoanalysis, which would dominate for the next fifty years, “insight” was defined as patient self-understanding of unconscious dynamics more broadly. Insight was indeed the sine qua non of treatment, for bringing the unconscious to conscious awareness was the mechanism of action in “insight oriented therapies.” With the majority of theory and practice devoted to non-institutionalized, wealthy, and “neurotic” voluntary outpatients, however, the age of psychoanalysis paid scant attention to the problem of psychotic illness awareness in the incarcerated mad population. Still, it is this body of work that provided a language of insight that could eventually serve that purpose, for psychoanalysis was the first variant of psychiatry that placed a heavy emphasis on the subjective thought processes of the patient.

By the 1970’s, psychiatric researchers found patients’ understanding of illness increasingly problematic and linked to a new object of interest beyond post-psychotic-depression. The first major publication in this vein, Van Putten et al. (1976), theorized the new phenomenon of outpatient drug refusal as the psychodynamic “wish to be crazy” (cf. Geisen and Feuer 1984 on the “need to stay crazy”). Such authors psychoanalyzed patients’ delusions of health as an “ego syntonic” coping strategy, with awareness and medication adherence viewed as serious threats to a grandiose sense of self. Consider this case description of a young man who thought himself a successful aeronautical engineer until he began taking antipsychotic medication.

While receiving fluphenazine decanoate, 25 mg every two weeks, he was objectively globally improved; subjectively, he seemed to resent the increased reality contact…He started to mention his loneliness and his realistic lack of any
life accomplishments and developed some insight into illness. At this point, he demanded to leave the hospital and resumed living in a lonely hotel. He returned for only one injection of fluphenazine decanoate and was readmitted three months later, psychotic as before (Van Putten et al. 1976, p. 1446).

Here the phenomenon of insight appears similar to Meyer-Gross’ work fifty years earlier—the patient’s growing awareness of madness threatened the self—yet the object of interest mediated by insight had changed. Upon developing insight, the young man encounters a depressing reality, but addresses this by simply demanding to leave the hospital and avoiding his outpatient injection follow-ups. Where Meyer-Gross (1920) was concerned with explaining why patients fell into post psychotic depressions, the researchers of the 1970-80’s were concerned with insight as the mediator of outpatient non-compliance. What accounts for this?

Put simply, psychiatry’s institutional needs had changed following the mid-twentieth-century restructuring of the Western welfare state, codified in commitment law and realized in wide-scale psychiatric patient deinstitutionalization. In the case of the United States, the burden of care was shifted from state governments to local communities and patients themselves, who were granted both new rights to self-determination and financial independence from caretakers in the form of new Federal welfare programs like SSI and Medicare (see Scull 1977 for a strong review of the historical process). Many caretakers were immediately skeptical of this self-determination, and as Van Putten states in the opening of an earlier article addressing drug side effects, “The reluctance of patients with schizophrenia to take their prescribed phenothiazines is the bane of the psychiatrist” (1974 p. 67). He cited recent urine colorimetric studies that found between 24%-63% of schizophrenic outpatients had taken less medication than prescribed, part of a growing sentiment that newly empowered patients were not choosing the proper course of action.

Where coercive treatment had been almost entirely within the discretion of mental health professionals, legislation such as California’s 1967 Lanterman-Petris-Short Act granted mental patients considerable new rights to treatment decision-making (see Scull 1988). With this legal empowerment and discharge of patients who would likely have been institutionalized in the past, outpatient treatment compliance became a site of clinical, political, as well as scientific interest. As McEvoy et al. (1981 p. 856) put it in an interview study of patient attitudes toward treatment,

The right to refuse medication is a legal right now being extended by federal courts to many voluntary and involuntary mental patients. However, little is known of the insight that chronically ill mental patients bring to the decision of whether or not to accept prescribed medication. (p. 856 emphasis added).

More so than Van Putten et al. five years prior, McEvoy et al. articulated that changes in the legal structure had brought forth a new problem area of patient management. By their logic insight and its lack had always existed, but researchers were only now attuned to it in the observable problem of non-compliance.

To put it in Hacking’s (1998) terms, the emerging challenges of underfunded community mental health care, the increasingly strict laws regarding involuntary commitment, and the rhetoric of patient choice created an ecological niche in which outpatient drug refusal seemed a major social problem and insight became its mediator. In relation to his niche conditions, the “wish to be crazy” was taxonomically legible.
given the earlier work on poor insight and post-psychotic-depression, as well as the psychoanalytic interest in disruptions to the ego. The crazed ex-mental patient “off his or her meds” was certainly visible on the streets and in media reports. Non-compliance behavior was culturally polarized as either moral failing of the sick role or a new type of anti-psychiatric political action. The psychodynamic rendering, furthermore, held polarized values like autonomy and rationality in tension, as a person could be volitional in their choice to remain irrational. The process of deinstitutionalization was itself polarizing, framed as either the extension of liberties or state abandonment of the vulnerable. Finally, the concept offered release, not so much to patients, but to family members and psychiatrists attempting to make sense of the new situation at hand.

Welfare state restructuring and the closing of the asylums, a reconfiguration of civil commitment law, and the pharmaceutical revolution in midcentury psychiatry were the institutional and material conditions of possibility for this changing ecological niche. The link between insight and compliance behavior appears self-evident today, but outpatient drug refusal only came to prominence with legally empowered patient self-governance in the community and the reliance on neuroleptics as a tool of management. Where previous interventions such as moral treatment or psychoanalysis involved an ongoing relationship between therapist and patient, antipsychotic medication offered the possibility of self-directed treatment. Urine colorimetric tests of medication levels in turn allowed for a quantification of adherence, materializing compliance behavior as tractable. Where Dekker recognized the free mad people that “know it not” in 1604, Hobbes and Locke simply assumed the mad could not know their own madness, and Mayer-Gross analyzed the institutionalized post-psychotic-depressive in the 1920’s asylum era, the self-governing, drug refusing psychiatric outpatient in denial of madness was born in the late 20th century.

Family members, with few resources from the state and suddenly stripped of formal power, were de facto tasked with managing their now legally empowered and (ostensibly) financially independent relations. Although some ex-patients smoothly transitioned into the community, either entering the workforce or accessing Federal benefits, many did not. The ensuing crises helped birth the family member social movement and powerful advocacy groups that brought attention to the failures of community mental health care, who argued that ex-patients were “dying with their rights on” (Treffert 1973). Critics of the new commitment laws like Appelbaum and Gutheil (1979 p. 313) invoked Van Putten et al. (1976) in legal discussion, writing “Psychotic patients may often have no unitary “will” as the law conceives it, but rather fluctuate back and forth between mutually exclusive desires, unable to resolve conflicting wishes.”

This conception of will as “conflicting wishes” was ill suited for these legal debates, however, and civil liberties groups defended patients’ rights to self-determination. Indeed, the rights of the mentally ill had become a foundational case on which ACLU theorists were addressing institutional “enclaves” where the Bill of Rights had yet to be respected (for a history see Walker 1990). Civil libertarian theorists like Morse (1982) argued specifically that the behavior of those deemed mentally ill was no less volitional than that of the sane, and that psychiatrists had no objective way of determining competency or dangerousness. Here is where it is more useful to think of Hacking’s vector of “cultural polarity” in terms of conflicting institutional logics. While we take for granted that health field is shaped by the logics of the market, the logic of the
democratic state, and the professional logic of care, this is a recent historical development (Scott et al. 2000). Rather than the grand problem of liberty in American “culture,” there was an actual historically specific battle between the medical welfare state’s logic of controlling care and the ACLU’s logic of expanding democracy into these untouched enclaves.

Although these niche conditions were ripe for the medicalization of insight as mediator of compliance, the research was not translated into the broader scientific, legal, or policy spheres. There are likely four interrelated reasons for this, two of which are directly in line with Morse’s (1982) comments: First, psychodynamic theory attributed significant agency and desire to mad people in their putative “wish to be crazy,” leaving it ill suited for translation into a legal paradigm that predicates capacity on volition. Second, insight had not been materialized into a measurable, scientific construct, and was therefore of no more use than the existing legal concepts like capacity or competence. Third, the civil libertarian activists were largely unchallenged, as the family member social movement had yet to cohere as a serious force to be reckoned with. Fourth, there was not a political opportunity to galvanize the public, such as the string of killings by putatively mentally ill individuals that served as a flashpoint in the late 2000’s.

Such counterfactual thinking is by definition speculative, but the logic will become clearer after the following empirical sections. I next demonstrate the role of measurement and representation in making insight a somatic phenomenon, rendering it translatable into the legal and policy fields. Although the “wish to be crazy” could not reconcile the conflicting institutional logics of democratization and paternalistic social control, a neurological variant offered new possibilities for reformulating the terms of the debate.

Making Insight Material: Rating Scales, Card Sorts, Brain Scans and Stimulation

In what follows I explain how poor insight came to be materialized, fitting it into a different medical taxonomy and enabling a translation into the legal and policy fields. Here I partially bracket the political question of liberal governance in order to show the trajectory and shaping of insight in the domain of psychiatric science. This is crucial, for it is precisely the autonomy of the scientific process from political concerns that gives it cultural power.

The somaticization of insight required social construction in the most literal sense, with scientists working to build material theories and representations out of an intuitive but problematic concept. Although compliance had been materialized through urine colorimetric testing, insight itself remained elusive. The institutional context of post-DSM III psychiatry exerted a series of new selection pressures, with diagnostic entities more likely to survive if they could be quantified and recast as tangible material objects. With the resurgence of biological psychiatry and the development of new measurement technologies, researchers saw a brain-centered conceptualization of insight as first possible and then scientific common sense. Insight scientists created scales to render classification liquid and transportable, utilized card sorting tests to localize cognitive functioning, and brain scans to produce visual representations of that localization, resulting in a theory of poor insight as analogous to the unawareness of partial paralysis found in brain-damaged stroke victims.
The statement from the DSM IV TR (1994) schizophrenia section reads,

A majority of individuals with schizophrenia have poor insight regarding the fact that they have a psychotic illness. Evidence suggests that poor insight is a manifestation of the illness rather than a coping strategy. It may be comparable to the lack of awareness of neurological deficits seen in stroke, termed anosognosia. (American Psychiatric Association 1994 p.304)

How had such a statement become possible? Anosognosia, coined by Babinski in 1917, typically referred to unawareness of hemiparesis or cortical blindness that could be linked to a localized brain lesion. The neurologist Hinterbuchner (1974 p. 590) had in fact offered poor insight into psychosis as a negative demarcation of anosognosia’s boundaries, stating “there is no justification for including under this term patients with no awareness of their cardiac disease or patients who are schizophrenic and have no insight concerning their illness.”

The first step in making insight material was the creation of validated research populations, for investigators could hardly administer neuropsychological testing or brain imaging until they had comparison groups of people with varying degrees of insight. Prior to the DSM III, psychiatrists typically theorized etiology with in-depth case readings of patients’ lives and thoughts, as in Van Putten et al.’s (1976) psychoanalytic explication of the young man’s “wish to be crazy.” Following sustained social scientific attack and new developments in somatic treatments, psychiatric science entered a legitimacy crisis and ensuing paradigm shift toward biological models (Luhrmann 2000).

The neo-Kraepelinians, borrowing heavily from the rating scale procedures of psychometrics, turned to statistical methods to address the ongoing problem of diagnostic reliability (Kirk and Kutchins 1992). Without biological tests for mental illnesses, investigators hunting for an underlying genetic, biochemical, or neurological etiology continued to use phenotypical and behaviorally oriented diagnoses with which to correlate somatic markers (Lakoff 2006). The types of standardization processes theorized by Timmermans and Berg (2004) were crucial for justifying the paradigm shift, and the “paper tools” (Klein 2001) of psychology, here the standardized and internally validated diagnostic scale, were central to psychiatric research. Like other diagnostic entities in this new institutional environment, insight would need to be made material in order to thrive.

The Present State Examination (Wing et al. 1974), used by the WHO for schizophrenia research and authors like Van Putten (1976), contained a small section with four questions regarding insight. In line with a general movement to advance the psychological disciplines, which produced standardized scales for increasingly refined dimensions of symptoms and sub-diagnoses, insight researchers began to create sophisticated means of measurement. In 1989 McEvoy et al. produced the Insight and Treatment Attitude Questionnaire (ITAQ), the first measure devoted solely to insight. Recall that McEvoy et al. (1981) had conducted an interview study on insight, but where that research was restricted to descriptive claims like “only 13 percent understood they were mentally ill” (1981 p. 856), the standardized scale enabled correlational findings like “…degree of insight was not consistently related to the severity of acute psychopathology, as measured on two structured scales” (1989 p. 43 emphasis added).
With the ITAQ, McEvoy et al. made insight into a quantitative object, with degrees that could be measured against yet more standardized scales.

David’s (1990) Schedule to Assess the Components of Insight (SAI) introduced a likert scale to capture finer grained dimensions, and conceptualized insight as consisting of three parts, namely recognition that one has a mental disorder, compliance with treatment, and the ability to re-label unusual events as pathological. Amador and Strauss (1990) developed what has become among the most popular scales in research, the Scale to Assess Unawareness of Mental Disorder (SUMD), which further distinguished between awareness of symptoms and attribution. This distinction allowed for a person to be aware of an ostensible symptom, yet capture the person’s “misattribution” of that experience to something besides mental disorder. Both the SAI and the SUMD were considered breakthroughs for examining insight as multidimensional rather than dichotomous, and providing more sophisticated measures to capture such dimensions.

A close reading of these scales is important for illustrating what information is elicited in particular constructions of insight. Anthropologically minded clinicians have criticized the scientific operationalization of insight for failing to take culture into account, with Johnson and Orell (1996) showing, for instance, that hospital discharge summaries in the UK more frequently label non-whites as lacking insight. Given the suggestion that poor insight may have been misdiagnosed in such settings because of insufficient cultural sensitivity, others (Tranulis et. al 2009, Roe et al. 2007) have advocated the development of alternative “narrative insight” measures. My concern, however, is tracing how particular interests become embedded through the operationalization and production of insight research, whether bio-psychiatric or culturally sensitive measures are at play. That the SAI explicitly frames compliance with treatment as a dimension, for instance, signals the relevance of the practical problem of outpatient drug refusal to conceptualizations of insight. Where a vulgar social constructionism might dismiss these scales as a case of informational “garbage in, garbage out,” I instead analyze what specific concerns are medicalized through the evaluation process.

The SUMD is worthy of particular attention because of its popularity, such that when insight it is said to be located in the frontal lobes, it is often insight as measured by the SUMD. In comparison with the four questions supplementing schizophrenia diagnosis in the Present State Examination, the SUMD has twenty multipart questions addressing such phenomena as “awareness of unusual eye contact” (Amador and Strauss 1990). The first item asks, “In the most general terms, does the subject believe that s/he has a mental disorder, psychiatric problem, emotional difficulty, etc?” (p.5). Three options, of “unaware,” “somewhat” and “aware” can be selected and incorporated into the score. Substantively, a vision of insight emerges that is consistent with David’s (1990) suggestion that compliance with treatment is a core component of insightfulness. Consider question 2, “awareness of the achieved effects of medication”:

What is the subject’s belief regarding the effects of medication? Does the subject believe that medications have lessened the intensity of his/her symptoms (i.e) if applicable?
0 Cannot be assessed. (Note: ALWAYS code a “0” on any item as MISSSSING DATA)
1 **Aware**: Subject clearly believes medications have lessened the intensity or frequency of his/her symptoms.

2

3 **Somewhat**: Is unsure whether medications have lessened the intensity or frequency of his/her symptoms, but can entertain the idea.

4

5 **Unaware**: Believes that medications have not lessened the intensity of frequency of his/her symptoms.

Thus awareness of illness, and therefore “insight,” becomes explicitly associated with medication compliance. Built into the instrument is the assumption that psychiatric medication is indeed unambiguously useful for all patients. The notion of ‘attribution,’ which follows the ‘awareness’ component, demonstrates the epistemic certainty built into the test. After item 5 questions whether “the subject is aware that s/he experiences delusions as such, i.e. as internally produced false beliefs,” 5b reads:

- How does the subject explain this experience(s)?
  - 0 Cannot be assessed/item not relevant
  - 1 **Correct**: Symptom is due to mental disorder.
  - 2
  - 3 **Partial**: Unsure, but can consider possibility that it is due to a mental disorder
  - 4
  - 5 **Incorrect**: Symptom is unrelated to a mental disorder.

Critics like Kirmayer et al. (2004) have been quick to point out that notions of “attribution” and “misattribution” are extremely loaded, for they presume that there is indeed a correct medical account, and Jones and Brown (2013) have further highlighted the political dimension of the insight construct as a naturalization of clinical authority. Sociologists of mental illness have long noted that those labeled mentally ill lose credibility in discussing that putative illness (see Goffman 1961, Smith 1978), but my goal here is not mere critique. What my analysis of the scales and ensuing correlational research highlights is precisely how researchers have operationalized this culturally embedded concept, not only naturalizing clinical authority but also setting the stage for materializing it as a tractable object. This is the standardization component of the materialization process, for these instruments helped create the research populations that allowed insight scientists to look for associations with various somatic measures.

In a landmark paper, Young et al (1993) administered the SUMD, the Wisconsin Card Sorting Test (WCST), the Verbal Fluency Test, and the Trail Making Test to 31 hospitalized patients diagnosed with schizophrenia. While neither the Verbal Fluency nor Trail Making tests were significantly correlated with the SUMD, performance on the WCST was predictive of insight measures. Given the WCST’s reputation as a “frontal lobe test,” and the failure of the other tests to show statistical significance, the researchers proposed that poor insight was likely localized to frontal lobes and “executive functions.” Young et al. began to draw the analogy with neurological disorders, framing alternative explanations of symptoms as “confabulations,” which are thought to be organic disturbances of the frontal lobes.
Of interest here is the observation that many explanations and misattributions elicited from the subjects with the greatest lack of awareness had a rigidly held, confabulatory-like quality to them. A number of studies (Kapur and Coughlan, 1980; Mercer et al., 1977, Shapiro et al., 1981) have come to view confabulation as an organically based frontally mediated inability to self-correct, the very attributes tested on the WCST and found to correlate with the SUMD scores.” (Young et al 1993 p.122)

This is what I mean when I say that, rather than a stable concept diffusing across domains, insight was changed in the measurement and materialization process. Consider how far this “lack of awareness” is from the “mad who know it not” or those with a psychodynamic “wish to be crazy.” The scales made the phenomenon tractable, compatible with other technologies and logics in the DSM revolution, and repurposed for new uses, despite traveling under a similar name.

Soon scientists turned to more expensive technological measurements in the comparison of differential insight populations, with structural brain imaging the most common way of identifying somatic correlates. As Dumit (2004) has shown, the “decade of the brain” saw a substantial monetary investment in imaging, with scientists researching all sorts of behaviors and thoughts via brain scans. This representation and visualization is a powerful component of the materialization process, for the phenomena becomes legible, if still underdetermined, to lay audiences. In the insight literature, a range of possible locations have been identified, at a level of granularity far smaller than the card sorting tests “frontal lobes” hypothesis. Although somatic correlates need not necessarily lead to medicalization, as in the “natural variation” framing of genetic correlates after the de-medicalization of homosexuality (Shostak et al. 2011), the institutional environment for insight was primed for pathologization.

In 1994 the DSM IV TR schizophrenia sections, co-chaired by insight researcher Xavier Amador, featured the analogy with anosognosia. The subsequent twenty years produced thousands of papers discussing the somatic correlates of poor insight. Consider this review of the literature in an FMRI study highlighting the role of the central midline structures (Bedford et al 2012):

…recent sophisticated structural imaging studies have revealed relationships with specific brain regions or with fronto-temporal white matter [13]. The majority have found evidence for a relationship between poorer insight and either volumetric reduction or thinning of various cortical midline regions, most towards the anterior (frontal lobe: [14] medial-orbital prefrontal cortex PFC: [15]; medial PFC: [16,17]; medial-superior PFC: [18]; anterior cingulate: [15,19]; paracentral lobule: [16]) but also the posterior parts (posterior cingulate: [12,19]; precuneus: [18,20]; [16]). Two studies found poorer insight to correlate with increased volume in anterior midline regions [21,22]. Finally, a variety of other regions of reduced volume have also been implicated in the studies reviewed above including dorsolateral prefrontal cortex (DLPFC [23]), insula [24] and temporo-parietal regions [20].

Note that poor insight is associated with both thinning and increased volume of cortical midline regions, but such controversy over where and with what effect are less important than the robust finding that insight (as measured by various scales) can be seen in the brain.
These sorts of brain mappings next enabled the possibility of somatic manipulations of poor insight, borrowed from neurological theory, which materialized insight in practice. This overlaps with the following section on translational interventions but it is useful here for demonstrating how insight was made physically tangible. Consider Levine et al. (2012), an experiment published in *Brain Stimulation* that took the DSM IV analogy between poor insight in psychosis and anosognosia for hemiparesis as a literal one, using a somatic treatment for stroke patients in psychiatric populations. The experiment attempted to improve insight among schizoaffective patients by placing cold water in their ears, as the use of cold caloric vestibular stimulation has been known to temporarily improve awareness of hemiparesis for anosognosiac stroke victims. The study produced ambiguous results, but more important than the outcome, the publication both highlights and furthers the development of insight as a real, somatically based object. Built from scales, localized by card sorts, and visualized in brain scans, scientists could now touch insight through the ear canal.

If this detour through the science of insight has appeared to move away from the liberal dilemma of madness, it is because I want to highlight the semi-autonomy of the scientific field. The Levine et al. (2012) ear experiment, for instance, was not justified in reference to problems of governance, but merely as an extension of interventions already found in neurological practice. The fact that these scientists speak the language of brain region rather than will is precisely what lends their discourse and practice legitimacy. In the following section I show how this work altered the field of possibilities in psychiatric practice, court deliberation, and social movement mobilization centered on the sociopolitical problem of psychiatric patient liberty.

**Translational Science: Interventions, Case Law, and Policy Frames**

Now rendered material, transportable, and institutionally validated by its inclusion in the DSM IV TR, advocates worked to show how impaired insight could be mobilized for practical use in addressing the dilemmas of liberal governance. In this section I focus on the translation of insight research into family mental health care, legal scholarship, case law, and policy debates over outpatient commitment and legal culpability, where it has been most thoroughly embraced. Here “translation” describes both how facts are stabilized and reconstructed across social fields, and used to enroll seemingly disparate actors in shared projects (Callon 1986, Latour 1987), and the movement of “translational science” that aims aim to bring findings to bear upon clinical or policy interventions.

What is striking in the case of insight is that bench science has not been translated into clinical, somatic interventions aimed at restoring compromised insight, as indicated in the above brain stimulation experiment, but primarily into behavioral therapies targeted at caretakers and policy prescriptions regarding commitment law. In this milieu the mad who “know it not” are not so much willful drug refusers with a “wish to be crazy,” but neurologically damaged subjects with deficits to be managed or remediated, like traumatic-brain-injury patients. Insight has been translated as a new neurological object with new connections, helping to constitute new coalitions and enable new forms of action. If the previous section primarily showed the way insight adapted to fit a new epistemic environment, this section deals with the niche construction and mutual fitting as insight goes on to engage the legal institutions that gave birth to it.
Given researchers’ failure to produce any significant somatic clinical interventions, insight discourse could have been ignored or died out. On the contrary, it has continued to grow, in large part due to social movement actors like the National Alliance on Mental Illness (NAMI) and the Treatment Advocacy Center (TAC), which have each adopted the anosognosia concept in their pamphlet literatures pertaining to involuntary commitment. Standing opposed to the ex-patient activists and ACLU, such organizations are composed primarily of family members and professionals. These groups sought to penetrate the institutional logics of psychiatry and the legal system, both by equipping lay actors with expert knowledge, and hiring experts to act as brokers (see Epstein 1996 on the case of HIV science). NAMI has become a particularly powerful advocate of bio-psychiatric logics, with critics calling attention to the fact that the pharmaceutical industry supplies a majority of its funding (Rothman et al. 2011).

Xavier Amador, the previously mentioned insight researcher, became a champion of the new knowledge, simultaneously acting as scientist, DSM knowledge gatekeeper, family advocate, expert witness, and moral entrepreneur. Amador was in fact closely aligned with the family member groups, identifying as the brother of a severely psychiatrically disabled man who was medication “non-compliant” (Amador 2000). After the publication of DSM IV TR he moved further outside of the academic realm, consulting in legal cases, writing popular texts, and offering commentary on news shows. By the early 2000’s he was leading NAMI’s Center on Practice and Research and later serving on the group’s Board of Directors.

Amador’s (2000) best-selling book for family members and clinicians, entitled “I’m Not Sick I Don’t Need Help!” is a key text for understanding the diverse manifestations of insight knowledge outside of the expert domain. In a telling anecdote of the book’s popularity, anthropologist Paul Brodwin (2013) reports outpatient psychiatric staff instructing him to read it prior to conducting fieldwork at their clinic. Conceding the intractability of anosognosia, Amador’s regimen is not a targeted intervention for the patient so much as a reorientation for caretakers, whose own irrationalities are subjected to analysis. Once they accept that they cannot convince patients of illness, caretakers are free to rationally assess what can be changed. Family members and therapists learn to better “partner” with the patient, reframe the effects of anosognosia in order to avoid burnout, engage with police during emergencies, and deal with the relational fall out of commitment proceedings. In contrast to the expected focus on normalization, Amador advocates a surprising tolerance of psychic difference and strange behavior. This is possible because patient subjectivity and resistance are, in his model, thoroughly medicalized.

Although some psychiatric service users have reacted to the insight concept with skepticism and indignation, others have taken it up as a resource for making sense of their lives. This appears paradoxical, for if anosognosia prevents the knowledge of mental illness, how can one then be knowledgeable about how mental illness impairs that knowledge? Yet as Novas and Rose (2000) have argued in their work on “somatic individualism,” people may respond to seemingly deterministic discourses of their somas not with resignation, but with new ethical practices for monitoring and regulating their
selves. Consider this (2012) blog post from a woman who identifies herself as suffering from impaired insight:

Anosognosia is the term for the most dangerous symptom of mental illness. It’s the belief that you’re not mentally ill and don’t need your meds…I know consciously that going off my meds would be a bad idea, but because of this symptom, the concept seems perfectly reasonable. Unlike many others, I have the sense to discuss my plan with my loved ones.

Note that the woman does not see her poor insight as a global impairment, but as a labile symptom that can be monitored. Her claim that she is equipped with enough “sense” to discuss this anti-medication plan with her loved ones, “unlike many others,” suggests an ethical responsibility of vigilance. She has come to see at least some of her free choices regarding treatment as potential symptoms, and her own volition as compromised and dangerous. The first response to the blog post nicely illustrates the consequences such a conceptualization has on selfhood:

Hello, a name for why I am delusional about medication …I am doing nothing, my kids are grown, my grandchildren don’t need me as their nanny anymore. All I ever do is humiliate them with my crazy bipolar Anosognosia. Stopping my medication to get “ME” back, focused, intelligent, organized and successful with relationships. (This person never really existed that is the delusion).

For this blogger, the impaired insight classification enables a new interpretation of her identity and kinship role. Although she says she previously stopped her medication to get her true self back, she now recognizes that this well-put-together-person was in fact a delusion. Such statements demonstrate the solace that some service users receive from the impaired insight classification, precisely the kind of “release” that Hacking (1998) discussed in his initial ecological niche conditions. Consider how the above bloggers might score on Amador’s SUMD test before and after exposure to the insight research, or receiving their own SUMD score. Hacking would likely trace the way these women’s altered practices and self-conception render the classification outdated and subject to revision in a “looping effect,” but here his account of feedback reaches its limit. In what follows, I illustrate a different form of feedback through institutions and knowledge infrastructures, rather than looping “through the head” of the classified.

Moving beyond clinicians, families and patients, legal scholars have deployed the impaired insight research to address two persistent points of contention in debates over involuntary commitment, such as those posed by Morse (1982). The first revolves around the question of how to ensure citizens maximum liberty, and the second involves the problem of subjectivity and error in competence evaluation. Responding to court decisions that view compulsory treatment as an impingement on liberty, legal scholar

---

Joanmarie Davoli (2002) utilizes the insight research to argue that it is in fact mental illness that threatens freedom, not coercive treatment. By preventing a person from appreciating his or her condition, the symptom of poor insight redefines any notion of free will. She writes,

Recent research reveals that denial or inability to accept that one is mentally ill is actually a symptom of the disease. Thus, an individual suffering from schizophrenia has a brain defect that is telling him that he is not sick and that he must refuse treatment. In this way, the disease itself actually steals the afflicted individual’s “free will” (Davoli 2002 p.1007).

Davoli’s article focuses on the importance of integrating such “recent research” into contemporary law, admonishing courts for relying on outdated psychiatric knowledge. The metaphor of an external agent that “steals” free will from an autonomous self is key here. Consider the difference between this metaphor and that of psychoanalytic denial theory, where an agent retains volition in his or her “wish to be crazy.”

Regarding the vagaries of evaluating volition, legal scholar Kenneth Kress (2004) further suggests that insight should replace the messy concept of competence as the standard by which courts assess capacity and justify involuntary commitment. This is because it had become objectively measurable. Given that it “might be more determinately and precisely measurable, it is plausible that in legal applications lack of insight will command more intersubjective agreement than competence,” and a material basis will “minimize decision-maker discretion, thereby reducing the risk of bias, prejudice, self-interest and the importation of subjective values” (Kress 2004 p.270). Here the insight scales are put to work as the material bedrock on which Kress’ argument rests, an objective way to evaluate the free will and volition that Davoli (2002) sees as central to American law.

Xavier Amador began translating his scientific expertise in a high profile, contentious and eventually aborted insanity defense. Anosognosia entered the criminal courtroom and public eye, not as an abstraction, but in the figure of “Unabomber” Theodore Kaczynski, on trial for killing three and injuring many more with homemade explosives. Famously refusing a state psychiatric evaluation, Kaczynski insisted that he had engaged in the attacks as part of his critique of technologized society, and the prosecution moved to use his non-compliance as a way to preclude any consideration of mental illness. Kaczynski’s attorneys, working with Amador, argued that this was actually anosognosia, and only further evidence of insanity. When they asked the judge to take impaired insight into consideration, Kaczynski moved to fire his defense and represent himself (see Amador and Shiva 2000). This request was denied and he ultimately accepted a plea bargain for life in prison.

Note that what is at stake is not Kaczynski’s criminal culpability, for that was not the target of Amador’s intervention, so much as the ability to reject the insanity plea in the first place. That Kaczynski might rather die than have his “manifesto” dismissed as insane was not considered, although eco-anarchists have indeed engaged the document as more than the ravings of a mad man (Noble 1995). Whether this makes Kaczynski the insightful one and society deluded is less important than the fact that impaired insight had circulated in both the courtroom and public discourse, with criticism soon following.

Philosopher-journalist Alston Chase (2000) denounced the defense’s logic in an extended piece in the Atlantic, writing, “a refusal to admit to being insane or to cooperate with
people who are paid to pronounce one insane cannot be taken seriously as proof of insanity.”

Successful insanity defenses are rare in America, and it is well established that somatic accounts of legal culpability are typically more relevant in sentencing than verdict (see Rosenberg 1968 for a classic 19th century case study, and Rose 2000 for contemporary analysis). Indeed, Amador’s expert testimony in other murder cases has typically failed to sway judges and juries (see for instance State v. Reid 2005). In civil cases, however, psychiatric expert witnesses have used the impaired insight concept as a potent force for assessing risk in cases to determine hospital release (see Keomany v. Commision of Human Services 2012) and the appropriate level of restrictive care (see Stuck v. Public Guardian 2008). Perhaps less obviously, experts and judges have also used the concept to complicate the meaning of compliance, reframing what might look like malingering as a potential confirmation of illness, and claims of illness and the need for treatment as proof of fakery. Although variations on the logic of insight and the “catch 22” of madness can be found in earlier case law, actors now point to anosognosia as a scientifically validated fact with significantly stronger rhetorical power.

Consider Collins v. Astrue, Dist. Court, WD Arkansas 2012, an appeal of a dismissed disability claim. The administrative law judge who first evaluated the case noted the claimant had failed to take prescribed medications for schizophrenia, indicating that he was not schizophrenic and profoundly disabled. The district court reversed this decision, invoking the DSM IV insight research to reach the opposite conclusion: the person in question had likely failed to take medication because of anosognosia for schizophrenia, therefore confirming mental illness and deservingness of disability benefits.

Consider next Hursey v. Taglia 2012, in which a state prisoner unsuccessfully brought civil rights charges against his psychiatric care providers. The plaintiff, Mr. Hursey, asserted that his antipsychotic prescription was discontinued as retribution for grievances he filed, that he was denied appropriate placement in a unit where he could receive care, and that this deliberate indifference to his medical need violated his Eighth Amendment rights. Defendant Taglia testified that Mr. Hursey’s insistence on his psychosis and need for medication confirmed that he was in fact a malingering, undermining both claims. The magistrate judge’s decision cites Taglia’s statement that Hursey “…actively asserts that he is psychotic, which is contrary to the anosognosia that is the hallmark of major mental illness.” Inverting the logic of the previous case, the knowledge of madness and desire for medication here confirm sanity and duplicity.

Although the precise impact of such information upon judges’ decision-making is difficult to discern, and judicial outcomes are further contingent on the organizational exigencies of particular courts, I use these two case examples to show how the impaired insight concept is deployed to make sense of contentious classificatory decisions in two primary dilemmas of madness in the neoliberal era, benefit provision through disability after the end of welfare-as-we-knew-it (see Dobransky 2009) and differential housing assignment in wildly overcrowded prisons (see Rhodes 2004). In each case it is widely acknowledged that there are “deserving” people with severe mental illness, but bureaucrats must still ration resources based on ultimately unstable psychiatric criteria. Rather than speculate on whether either plaintiff was in fact malingering or severely psychiatrically disabled, I point instead to the fact that impaired insight here acts as a now
materialized and institutionally legitimated conceptual resource for both affirming and denying deservingness.

What I address next is the way the anosognosia concept has moved beyond the confines of individual court cases to address nearly all actions that a psychiatrically disabled person might take. The insight research has increasingly been mobilized to explain larger and larger social problems, tying macro-level phenomena and potential policy solutions to the micro-level insight/compliance nexus. Here we can observe a significant function creep, with impaired insight seemingly accounting for everything that goes wrong in community mental health care. Consider Lamb’s (2009) editorial in the American Journal of Psychiatry on the failure of deinstitutionalization and the current criminalization of mental illness:

A large proportion of persons with severe mental illness who have committed criminal offenses and are now in jails or prisons are found to be highly resistant to psychiatric treatment. They may refuse referral, miss appointments, fail to adhere to their medication regimen, engage in substance abuse, and refuse appropriate housing placements. Many of these persons suffer from anosognosia, a biologically based inability to recognize that one has a mental illness—and thus a biologically based lack of insight.” (Lamb 2009 p.9)

A close reading of this editorial is revealing of the ease with which lab based research is extended to explain an increasingly vast array of phenomena. Lamb, a respected psychiatric leader and member of NAMI’s board of directors, embraces the anosognosia analogy wholeheartedly, without the qualifiers found in the DSM IV TR phrasing, and moves beyond medication non-compliance to link a host of behaviors, interactions, and outcomes, to impaired insight. Even a complex phenomena like refusal of “appropriate” housing placements, often attributed to dissatisfaction with ghettoized housing stock or paternalistic treatment requirements (Tsemberis et al. 2004), is tied to anosognosia. Lamb concludes by suggesting that the best way to reverse the criminalization of mental illness is to expand involuntary treatment options like assisted outpatient treatment (AOT), or coercive care outside of the hospital.

The AOT question became national news in the wake of the Sandy Hook elementary school shootings, where a young man presumed to be mentally ill gunned down twenty-six people. Less than a year before the Sandy Hook killings, the ACLU defeated proposed AOT legislation that would facilitate outpatient commitment in Connecticut. Proponents have suggested that Adam Lanza, the ostensibly mentally ill perpetrator, might have been identified and detained had the AOT law gone into effect. Although various forms of outpatient commitment have been experimented with since the 1980’s, and AOT is on the books in 42 states, only 18 have actually operating programs. The Courant, Connecticut’s largest newspaper, hosted an op-ed debate addressing insight and AOT between community psychiatry professor Larry Davidson, known for his research on “recovery” and the patient led advocacy movement, and Xavier Amador.

Davidson (2013) argues that defining insight is inherently problematic, and that the scientific evidence base does not exist. He writes, “There are no data of which I am aware that indicate that persons with psychotic disorders refuse treatment because they have a neurological condition that makes them unaware that they are ill.” Instead, Davidson suggests that making effective mental health care for people with mental illnesses requires “fixing the society, and system, that marginalizes them.” Here he
echoes a common trope in progressive mental health circles, locating treatment non-compliance in a social system that excludes, oppresses, and ultimately fails psychiatric patients. Like Farmer’s (1998) notion of structural violence, he suggests that mental patients often struggle to remain in treatment because of inadequate social and material support.

Amador’s (2013) reply is simple and rhetorically powerful, noting the extensive anosognosia research and his own scientific bona fides to call Davidson’s expertise into question. He writes, “I would point him [Davidson] to page 304 of the DSM IV which summarizes the anosognosia research gathered by a committee I co-chaired…” By invoking his own authority as a co-chair of the DSM IV text revisions and the authority of the diagnostic manual itself, he dismisses Davidson’s concerns as simple ignorance of the science. He then writes, “Anosognosia is real and AOT provides a workaround that can save lives.” Note that, despite their differences, the debaters agree that patients themselves are not the locus of non-compliance. Where Davidson sees the failure to address a broken social system, Amador sees the failure to properly administer compulsory treatment, but neither invokes a volitional “wish to be crazy.”

Finally, we turn to the actual legislative attempts to overhaul commitment law, framed and predicated on the slow materialization and translation of insight. In 2013 Representative Timothy Murphy of Pennsylvania introduced the “Helping Families in Mental Health Crisis Act,” H.R. 3717, which aims at a massive reprioritizing of mental health dollars for the national expansion of AOT. Murphy, a psychologist by training, justified this coercive care in reference to research on impaired insight for psychosis. In his speech to the House floor he explained,

We know fifty percent of persons with schizophrenia suffer from a neurological impairment that makes them incapable of understanding they are ill. This lack of awareness, termed Anosognosia, is the leading cause of noncompliance with psychiatric treatment. This neurological problem helps to explain why forty percent of Americans with a serious mental illness do not receive treatment, and it explains how our system fails to help those most in need. (CSPAN 2013)

Consider next Maryland House Bill 767 (2014), a proposal for a new AOT program. Two clauses from the preamble are worth quoting at length:

WHEREAS, When individuals with anosognosia remain untreated, they suffer unnecessarily from homelessness, poverty, repeated arrests, trauma, and suicide; and

WHEREAS, When individuals with anosognosia remain untreated, they lack the capacity to make rational and informed decisions regarding their civil rights…

Note that civil rights here are not denied by forcible treatment, but by mental illness itself, much like the legal scholar Davoli’s (2001) conceptualizations presented above. Further, HR 3717 frames “how our system fails” and the Maryland bill explains patient homelessness, poverty, and adverse interaction with law enforcement in terms of impaired insight, rather than structural explanations, akin to Lamb’s (2009) above claims. Although is difficult to evaluate the impact of these legislative attempts, as H.R. 3717 remains under House review, and the authors of Maryland HB 767 are regrouping after a recent defeat, such legislative attempts demonstrate the role of the impaired insight concept in debates over mental health policy and patient rights.
Recall that in 1979 Appfelbaum and Gutheil invoked insight research to argue against the newly libertarian changes to commitment law. They claimed that patients’ “conflicting wishes” were far from the “unitary will” required for defining a person as volitional in the legal sense, but such writings, and the claim that discharged patients were “dying with their rights on,” failed to generate serious challenges to the new commitment standards. I suggested that this was due to incompatibility between the volitional “wish to be crazy” theory and legal conceptions of free action, the problem of vagueness in measurement, the weakness of the burgeoning family member social movement vis-à-vis the civil libertarians, and a lack of political opportunity.

Today, legal scholars take the logic of brain-based poor insight as compromising volition, scientists use scales and scans to render anosognosia measurable, the family member social movement is a powerful lobby with pharmaceutical industry backing, and a series of national tragedies involving disturbed individuals and gun violence have created a flashpoint political opportunity for commitment reform. Here with the material and political conditions in place, the importance of impaired insight research in commitment law reform is not only thinkable, but a powerful and actionable frame. From a folk concept, to a psychodynamic theory, to a measured psychometric object, to a neurological pathology, to a legal strategy, and finally a micro-medical-frame for addressing macro-social problems like homelessness and trans-institutionalization, insight has not only transformed in relation to changing conditions, but also helped facilitate action that altered the conditions in question.

Discussion

To summarize the findings in relation to the theoretical apparatus: First, deinstitutionalization and the legal empowerment of psychiatric patients created an ecological niche in which patient self-knowledge became tied to the problems of non-compliance. The “wish to be crazy” addressed the conflicting institutional logics of social control and democratization, but gained limited traction in the larger institutional environment because of a disjuncture between psychodynamic thought and legal understandings of will. Changes in that environment, namely the paradigm-shift in psychiatry and a growing interest in commitment law reform, created a series of selection pressures, and with the introduction of new representative technologies, the wish to be crazy was materialized as impaired insight or anosognosia. Following this adaptation, impaired insight enabled new forms of legal thought and argument, facilitating its translation into legal scholarship, actual court cases, and current social movement and legislative attempts to address the very questions of custodial management that birthed it. This, I have argued, is fruitfully theorized as a form of classificatory niche construction.

To demonstrate the intellectual “cash value” (James 1907) of my niche construction approach, consider the limitations of existing theories of power and classification in the human sciences. Labeling theorists might alert us to the way the impaired insight classification medicalizes patient dissent and cultural variance, but the focus on interaction and power in the present would have done little to illuminate the historical dimensions of emergence. Hacking’s dynamic nominalism helps explain the ecological niche of the “wish to be crazy” in 1970’s deinstitutionalization, and the potential “looping effect” in the 2000’s, but the focus on emergence and then looping
through the minds of classified people ignores the complex feedback and transformation of insight through scientific and legal institutions. By offering a unidirectional account of ecological dynamics, and by skewing towards human agency in his account of feedback loops, even Hacking’s sophisticated approach fails to see the ways classifications remake the surrounding ecology.

I have shown that impaired insight was not merely used by pre-existing actors for pre-existing goals—it enabled new coalitions of actors across social fields, forms of legal critique, changes in knowledge infrastructures, and techniques of self-construction. To account for the dynamic way the classification does things when translated across social fields, we must move outside of the heads of the classified and into the material world of scales, brain scans, legal documents, case law, and practices. This is what is gained by building on Hacking’s ecological niche with adaptation and niche construction processes, or the mutual fitting between entities and environments. In order to get at these processes, I have advocated an expansion of Hacking’s empirical scope both “down” to technical questions of scientific practice and “up” to institutional analysis.

**Conclusion:**

The dilemmas of psychiatric patient self-determination are, I have argued, specifically liberal ones. In an authoritarian state, where individual freedoms are little valued, the confinement of mad people for their own and society’s protection is unproblematic. In a truly libertarian government, where individual freedoms are privileged (at least until people violate criminal law), paternalistic “care” is scorned for intruding on personal liberty. Recall that thinkers like Thomas Szasz (1961) and Morse (1982) went beyond the classical liberals in advocating criminal trial as a fairer form of judgment than hospitalization without due process. Yet as much as the American criminalization and social abandonment of mad people appears to resemble this model, the question of “good despotism” remains. Rather than stake a claim as to the proper resolution of these liberal dilemmas, or comment on the “reality” of impaired insight, I have attempted to historicize the terms of these debates and show the contingency of the logics framing them as moral and scientific positions.

The case presented speaks to numerous other situations where medicalization logics address the tension between paternalistic governance and individual rights. The insanity plea, which I briefly touched upon, hinges on the same questions of volition and whether scientific expertise can or should penetrate legal practice (see Erickson and Erickson 2008). Defendants seeking an insanity defense must prove their own madness and compromised volition, inverting the problem in civil cases, but with different standards of classification tied to different institutional trajectories in the criminal law. Street-level-bureaucrats encounter these dilemmas even prior to the courts, as in the informal diversionary programs of “therapeutic policing” that diagnose and process certain homeless people as appropriate for mental health and addiction services rather than arrest (Stuart forthcoming).

Nor are attempts to scientifically resolve the tension of paternalism and liberty restricted to the mad or deviant. Consider Obama advisor Cass Sunstein’s (2014) claim that the discoveries of behavioral economics undermine the Millian dictum that individuals are best positioned to make their own decisions. This is because “Mill’s view
neglects the existence of behavioral market failures, and the wide range of behavioral findings about human errors” (2014 p. 81-82). Indeed, the entire field of behavioral economics can be understood as an attempt to scientifically understand why people do not act “in their own best interest” and how to make them do so, a particular problem for liberal states that seek to govern while ostensibly maximizing freedom. Given bounded rationality, Thaler and Sunstein (2008) assert that even the sanest of citizens will benefit from “choice architects” who can “nudge” them into the proper prudent and healthy decisions, with neuroscience and behavioral economics justifying a new politics of “libertarian paternalism.”

Moving beyond the dilemmas of liberal governance, *classificatory niche construction* should prove useful for studies of classification and power more generally. Building on labeling theory and Hacking’s dynamic nominalism, my approach directs attention to the material, legal and infrastructural consequences of a dynamic “mutual fitting” between classifications and environments that has been largely ignored. If labeling theory alerted sociologists to the power of reified classifications to influence treatment of individuals as well as their self-conception, and Hacking opened the analysis to historical emergence and feedback beyond the self-fulfilling prophecy, my theoretical apparatus looks outside of human beings, to the way classifications shape the knowledge infrastructures, material objects, and institutionalized practices with which we think, act, and are processed.

**References**


H.R. 3717 (113th Congress) “Helping Families in Mental Health Crisis Act.”


Kraepelin, Emil. 1917. Lectures on Clinical Psychiatry, by Dr. Emil Kraepelin... Wood.


State v. Reid, 213 S.W.3d 792 (Tenn. 2006).


