Integrating Insults: Using Fault Tree Analysis to Guide Schizophrenia Research across Levels of Analysis

Angus W. MacDonald III 1*, Jennifer L. Zick 2, Matthew V. Chafee 2,3 and Theoden I. Netoff 4

1 Department of Psychology, Translational Research in Cognitive and Affective Mechanisms, University of Minnesota, Minneapolis, MN, USA, 2 Department of Neuroscience, University of Minnesota School of Medicine, Minneapolis, MN, USA, 3 Veterans Affairs Medical Center, Minneapolis, MN, USA, 4 Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN, USA

The grand challenges of schizophrenia research are linking the causes of the disorder to its symptoms and finding ways to overcome those symptoms. We argue that the field will be unable to address these challenges within psychiatry's standard neo-Kraepelinian (DSM) perspective. At the same time the current corrective, based in molecular genetics and cognitive neuroscience, is also likely to flounder due to its neglect for psychiatry's syndromal structure. We suggest adopting a new approach long used in reliability engineering, which also serves as a synthesis of these approaches. This approach, known as fault tree analysis, can be combined with extant neuroscientific data collection and computational modeling efforts to uncover the causal structures underlying the cognitive and affective failures in people with schizophrenia as well as other complex psychiatric phenomena. By making explicit how causes combine from basic faults to downstream failures, this approach makes affordances for: (1) causes that are neither necessary nor sufficient in and of themselves; (2) within-diagnosis heterogeneity; and (3) between diagnosis co-morbidity.

Keywords: reliability engineering, fault tree analysis, schizophrenia, DSM-5, research domain criteria, psychosis, NMDA receptor

The current framework for categorizing psychiatric disorders such as schizophrenia impedes progress toward the goals of understanding the causes and cures for these illnesses. This article will describe an alternative framework derived from reliability engineering, a field developed to study the way in which complex systems break down. Changing a framework is more than window dressing: frameworks carry the assumptions and affordances that invisibly guide our work. We will first critique two prominent frameworks in psychiatry. We will then show how a framework inspired by reliability engineering can help understand the cognitive and affective dysfunctions of psychosis.

The neo-Kraepelinian framework is named for the founder of modern psychiatry (Kraepelin and Diefendorf, 1907), built on schemes such as the Feigner criteria (Feighner et al., 1972) and the Research Diagnostic Criteria (RDC; Spitzer et al., 1975), and was eventually codified in the 3rd edition of the DSM (DSM-III). These codes used symptoms to determine whether someone fulfilled the necessary and sufficient conditions for a diagnosis, which assumed the status of a natural category. Although this made it possible for patients to share a diagnosis without sharing any symptoms, this level of within-category variation was acceptable if the criteria increased diagnostic reliability and harmonized practice (but see Markon et al., 2011). Within this framework theories about how the neural functions of schizophrenia patients are distinct from the neural functions...
of depressed or bipolar patients were immediately salient and substantive. Four decades later, genetic, cellular, neural, cognitive and affective dysfunctions are known to be shared across many distinct diagnostic categories, and even an optimal neo-Kraepelinian scheme could not sort things out. Fifty percent of people diagnosed with schizophrenia will fulfill criteria for comorbid substance abuse, and 50% will fulfill criteria for depression (Buckley et al., 2009; for additional critique, see MacDonald, 2013). Such comorbidity likely derives from a shared vulnerability: the genetic correlation for common SNP's between schizophrenia and bipolar disorder may be as high as 0.68, and between schizophrenia and depression may be as high as 0.47 (Lee et al., 2013). At this point the neo-Kraepelinian framework helps insurance adjusters more than researchers, clinicians or patients. For these reasons, there is momentum toward another framework that we will call informal reverse-engineering.

Informal reverse-engineering, already implicit in much psychopathological research, is codified in NIMH’s Research Domain Criteria (RDoC) program: “The mandate for RDoC is to consider psychopathology in terms of maladaptive extremes along a continuum of normal functioning, to promote a translational emphasis” (Ford et al., 2014, p. S296). The framework of RDoC is a matrix: rows are different functions from five categories (positive valence systems, negative valence systems, cognition, social processes and arousal); columns are levels of analysis ranging from genes to behavior and symptoms (Insel and Cuthbert, 2009; Sanislow et al., 2010; Cuthbert and Kozak, 2013; Ford et al., 2014). Patients can therefore be characterized in terms of many different functions and how those functions relate to biological processes or broader symptoms and social dysfunctions. Moving the focus toward functional deficits and away from the clinical constellations of the DSM may enable a tighter link between biology and behavior in neuropsychiatric research. However, a number of concerns have already arisen. First, by its structure, RDoC implies independence across the brains’ multitude of functions (rows in the RDoC matrix). Just as DSM categories are comorbid, so different RDoC functions are often correlated (reminiscent of general cognitive ability; see also Fowler et al., 2012). The authors of RDoC are aware of this, but there is little allowance in the framework for elaborating these relationships. Second, the RDoC posits that variation in these functions improves our ability to identify biological mechanisms more than clinical symptoms can. However, such functions may turn out to be as complicated as symptoms or diagnoses (Flint and Munafò, 2007; Kendler and Neale, 2010). Third, the relationship between these functions and symptoms is weak at best (Gold et al., 2012), and at times functions associated with very different neural structures correlate equally well with symptoms (MacDonald, 2013). The key shortcoming of the informal reverse-engineering approach in RDoC derives from isolating constructs from each other and thereby drawing attention away from the structure of psychopathology. Even when statistical relationships are discerned, it is not easy to append these into the cumulative science of mental illness.

This article will argue for a new framework, borrowed from reliability engineering, to guide the accumulation of knowledge and the development of new treatments for mental disorders. Our focus will be on psychosis, but many of these observations apply broadly to psychopathology. We will argue that adjusting the framework from which we motivate and report our findings is not only desirable, but is also a necessary step in the cognitive and affective neuroscience of these disorders. This new perspective based on a tool called fault tree analysis makes affordances for examining: (1) how causes that are neither necessary nor sufficient in and of themselves can result in psychosis; and how (2) within-diagnosis heterogeneity; and (3) co-morbidity arise.

**FAULT TREE ANALYSIS**

Fault tree analysis is a tool adapted from reliability engineering, in which the total likelihood of failure of a system is explained by the failure likelihood of each of its components (Rausand and Hoyland, 2004). Originally developed for analysis of rocket failures, fault trees seek to illustrate how faults of individual components interact within the system to cause an overall failure. To generate a fault tree, the different elements of the device must be identified, as well as their probability of failure—called faults—and how the faults interact and combine into failure modes (see Figure 1). A fault means the component is unable to perform its required function. Depending on other settings within the system, a small fault may be insignificant or trigger a cascade leading to a failure mode. In psychiatry, a failure mode is an impairment in some aspect of cognition, emotion or behavior of the kind we identify as a mental disorder.

The failure of a single component rarely results in a general system failure, due to built-in redundancy and plasticity of the brain. This enables most people with some forms of insult to function normally in the world. In this way, fault tree analysis makes affordances for causes that are neither necessary nor sufficient to cause dysfunction. Different individuals may come to a failure mode based on a different set of conditions, which contributes to within-diagnosis heterogeneity. Similarly, a given fault may lead to several failure modes, which gives rise to co-morbidity.

To generate a fault tree, one identifies the different components that contribute to the failure and for added value, quantify the likelihood of failure. Importantly, the fault tree illustrates how faults interact. This can be done with something like Boolean logic; faults may be related to other faults through AND or OR relationships (as well as others). This allows one to predict the probability of a failure mode given specific failures within the system. To this end, genetic mutations leading to psychiatric diseases may have measurable fault rates, although developmental and physiological failure rates will require more attention to estimate accurately. In addition, treatments can be modeled as externally-controlled variables in this network used to alter the outcome. The validity of the fault tree can then be tested by determining how well the known risks predict the known characteristics and rates of failure within a population.
Fault tree analysis is therefore a distinct and unused tool for understanding how different risk factors and therapies interact. Whereas the neo-Kraepelinian framework focuses on validating diagnoses and the informal reverse-engineering framework focuses on translational accounts of specific cognitive constructs, the reliability engineering framework attends to the ways in which faults combine into failure modes, and the ways in which the natural redundancies built into the brain can break down, or be built up. In this effort, computational models have a particularly important role.

**Computational Models as the Translational Glue within Fault Tree Analyses**

In many respects computational models—mathematical formalizations of hypotheses—are already providing the kind of mechanistic understanding of mental disorders such as psychosis that fault tree analysis promises. Across all levels of analysis, computational models provide a means for describing how a system will respond to perturbations (for review, see Rolls et al., 2008). Such perturbations might be thought of as faults—specific malfunctions that either do or do not lead to a failure mode—or treatments intended to reverse those faults. Integrating computational modeling with fault-tree analysis in the context of psychiatric disease allows us to bridge from risk factors, whose rates serve as the parameters of the model, to disease pathophysiology, which form the modeled outcome. Such models suggest links between different faults, and how those faults at one scale of the fault tree lead to outcomes at another scale.

For example, genomic studies have identified many mutations related to schizophrenia, based on weak correlations in very large samples. These are not direct links from cause to behavioral outcome. More than 100 risk genes have been identified by GWAS linkage studies (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), and the final number of relevant mutations may be thousands. Computational models relate what we may know about the functions of these genes...
Generating a Fault Tree: An Example in Schizophrenia at the Cellular Level

The work of Neymotin et al. (2011) can illustrate how a fault tree may be generated from the results of empirical research. The goal of this study was to simulate the effects of administration of the schizomimetic drug ketamine on hippocampal neuronal oscillations. Ketamine, an NMDA receptor (NMDAR) antagonist, elicits an increase in gamma (30–100 Hz) power and decrease in theta (3–12 Hz) power. The model consisted of 200 each of basket and oriens-lacunosum moleculare (OLM) interneurons and 800 hippocampal pyramidal cells and allowed them to test whether selectively blocking inhibitory circuits led to more high-frequency activity (Greene, 2001). NMDARs were present on the somas of each of the three cell types, as well as on the apical dendrites of pyramidal cells. Blocking NMDARs at all four locations decreased both theta and gamma power, which was inconsistent with ketamine effects in biological systems. A systematic investigation of all 16 possible combinations of insults demonstrated that pyramidal somatic NMDARs were largely irrelevant to theta and gamma power, whereas blocking pyramidal apical NMDARs alone was enough to decrease both theta and gamma power. The only condition which resulted in decreased theta and increased gamma was when OLM receptors were blocked while the basket interneurons and apical pyramidal receptors remained functional. This result can be translated into the Boolean logic illustrated in Figure 2.

The finding illustrates a helpful way of thinking: how do different insults combine to link NMDAR dysfunction to impairments in the cognitive and emotional functions of patients with schizophrenia? It may be that the cellular level of analysis is particularly amenable to fault tree analysis. Fault-tree models of cell-level malfunctions test causal relations between genetic mutations and changes in cell and network physiology, and then link these physiological changes to changes in cognitive and perceptual functions. Characterizing changes in brain function at the cellular level will require the development of appropriate animal models of psychiatric disease. Mutant mouse models link genetic mutations that increase risk for schizophrenia in humans to downstream changes in neural and network function. However, mouse models will be limited in the extent to which they can replicate cognitive and perceptual changes that occur in patients with schizophrenia, or the physiological dynamics of regions like the prefrontal cortex. Although it is difficult to relate cellular physiology and cortical network dynamics to downstream manifestations such as clinical symptoms, animal models can play an important role here. For example, ketamine administered to monkeys performing a cognitive control task show the specific pattern of errors that is a hallmark of cognitive control failure in patients with schizophrenia (Blackman et al., 2013). Neural recording during the period of synaptic malfunction also shows the underlying changes in cellular physiology and network oscillations that cause cognitive failures much like those seen in patients (Wang et al., 2013; Ma et al., 2015). Building links between synaptic, cell, and network phenomena in the context of behavior should serve to fill in key elements of fault tree networks.

There are other aspects of fault tree analysis that are not illustrated in Figure 2, for example the potential for two or more different combinations of insults to produce the same downstream effect, and modeling of non-binary outcomes using more complicated algorithms. A more complex model with more types of cells and insults may have revealed several mechanisms by which a decrease in theta power and increase in gamma power may result.

A FAULT TREE-ALIGNED RESEARCH AGENDA: IMPLICATIONS FOR THE COGNITIVE AND EMOTIONAL DYSFUNCTIONS OF PSYCHOSIS

It has been argued that schizophrenia is a syndrome not a disease: a set of symptoms, and other measurable signs, often seen together (Kotov et al., 2011; MacDonald, 2013). Assemblies of these failures, based on their relative co-occurrence, has led to our current diagnostic neo-Kraepelinian scheme. However, two patients labeled schizophrenic can both fulfill the current criteria without sharing any symptoms while simultaneously resembling patients not classified as schizophrenic. From a neo-Kraepelinian perspective, this is troubling. There is a clear need for a new framework that diagnoses patients in a new way.

Syndromal diseases are caused by multiple upstream faults resulting in a spectrum of symptoms, and no single cause defines the disease. The fault tree maps the likelihood of symptoms to the likelihood of cognitive and affective failures and in turn linking those failures to the likelihood of upstream faults in neural systems and genetic polymorphisms. This perspective allows for
the overlap of symptoms seen across syndromes. The fault tree’s logic gates relate the different components of the genomics and physiology into a biometric. While the goal is to generate an overall model that explains the direct mechanistic connections underlying the indirect correlations, this can be approached one step at a time. Parts of a fault tree can be assembled from existing data, perhaps using probabilistic graphical models, to link different elements at adjacent scales and correlational relationships between elements at nonadjacent scales. For example, linking genetic mutations to cell pathophysiology, or cell excitability to network oscillations. But links also occur between scales that are not directly coupled, such as correlations between genes and disease prevalence. The fault tree can bring these lines of evidence together to make predictions for interventions that can be tested with computational models and validated in animal models and clinical studies. In this regard, studies of the neural systems of the cognitive and emotional dysfunctions in people with psychosis, including schizophrenia, will play a critical role.

There are limitations to the fault tree metaphor. Physiology may not fall into a neat tree-like structure. A tree implies that there is directionality in the causality: genes are responsible for cellular physiology and physiology is responsible for behavior. However, in biological systems there is clearly feedback between each of these levels, which will require attention. Furthermore, the schema presented here relies exclusively on dichotomous inputs and outputs. For the most part variables in neuroscience are measurable quantities, frequencies or probabilities. This concern has been considered and has been addressed within the field of reliability engineering. The expansion of the Boolean fault tree into one with more realistic kinds of variables is an additional refinement. Studies of patients can aid in this effort by providing more data about the distributions of their measured variables. For example, the relative skew of patient and control distributions for neural variables is relatively unknown; such data provide important additional information for validly predicting the likelihood of rare events. Another challenge is that fault tree analyses are top-down, and therefore, even with computational modeling, require a good deal of prior knowledge. One way to address this challenge in large data sets with insufficient prior knowledge is the use of probabilistic graphical models, which can bootstrap these efforts from the bottom-up.

The ultimate goal in developing a new diagnostic schema is to identify combination therapies that will maximize benefits. Fault trees can help identify how genetic, pharmacologic or behavioral therapies can improve, and inform cost/benefit analyses when considering treatment options. This approach may also inform public health policy, helping to identify disease mechanisms for which intervention has not yet been fully exploited. A fault tree provides a formal way of relating disparate sources of
information about disease mechanisms and symptoms into one framework. Most importantly, it provides a path to move beyond a diagnostic scheme to one that directly therapies and research to maximize patient benefit.

REFERENCES

Blackman, R. K., MacDonald, A. W., and Chafee, M. V. (2013). Effects of ketamine on context-processing performance in monkeys: a new animal model of cognitive deficits in schizophrenia. *Neuropsychopharmacology* 38, 2090–2010. doi: 10.1038/npp.2013.118

Buckley, P. P., Miller, B. J., Lehrer, D. S., and Castle, D. J. (2009). Psychiatric comorbidities and schizophrenia. *Schizophr. Bull.* 35, 383–402. doi: 10.1093/schbul/sbn135

Cuthbert, B. N., and Koiz, M. J. (2013). Constructing constructs for psychopathology: the NIMH research domain criteria. *J. Abnorm. Psychol.* 122, 928–937. doi: 10.1037/a0034028

Ehrlichman, R. C., Satal, M. J., Maxwell, C. R., Lazarewicz, M. T., Finkel, L. H., Contreras, D., et al. (2009). N-methyl-D-aspartic acid receptor antagonist-induced frequency oscillations in mice recreate pattern of electrophysiological deficits in schizophrenia. *Neuroscience* 158, 705–712. doi: 10.1016/j.neuroscience.2008.10.031

Feighner, J. P., Robbins, E., Guze, S. B., Woodruff, R. A., Winokur, G., and Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Arch. Gen. Psychiatry* 26, 57–63. doi: 10.1001/archpsyc.1972.01750190059011

Flint, J., and Munafò, M. R. (2007). The endophenotype concept in psychiatric genetics. *Psychol. Med.* 37, 163–180. doi: 10.1017/s0033291706008750

Ford, J. M., Morris, S. E., Hoffman, R. E., Sommer, I., Waters, F., McCarthy-Jones, S., et al. (2014). Studying hallucinations within the NIMH RDoC framework. *Schizophr. Bull.* 40, S295–S304. doi: 10.1093/schbul/sbu011

Fowler, T., Zammit, S., Owen, M. J., and Rasmussen, F. (2012). A population-based study of shared genetic variation in preclinical IQ and psychosis among male twin pairs and sibling pairs from Sweden. *Arch. Gen. Psychiatry* 69, 460–466. doi: 10.1001/archgenpsychiatry.2011.1370

Gold, J. M., Barch, D. M., Carter, C. S., Dakin, S., Luck, S. J., MacDonald, A. W. I., et al. (2012). Clinical, functional and intertask correlations of measures developed by the cognitive neuroscience test reliability and clinical applications for schizophrenia cohort. *Schizophr. Bull.* 38, 144–152. doi: 10.1093/schbul/sbr142

Greene, R. (2001). Circuit analysis of nmdar hypofunction in the hippocampus, in vitro, and psychosis of schizophrenia. *Hippocampus* 11, 569–577. doi: 10.1002/hipo.1072

Hong, L. E., Summerfield, A., Buchanan, R. W., O’Donnell, P., Thaker, G. K., Weiler, M. A., et al. (2010). Gamma and delta neural oscillations and association with clinical symptoms under subanesthetic ketamine. *Neuropsychopharmacology* 35, 632–640. doi: 10.1038/npp.2009.168

Insel, T. R., and Cuthbert, B. N. (2009). Endophenotypes: bridging genomic complexity and disorder heterogeneity. *Biol. Psychiatry* 66, 988–989. doi: 10.1016/j.biopsych.2009.10.008

Kendler, K. S., and Neale, M. C. (2010). Endophenotype: a conceptual analysis. *Mol. Psychiatry* 15, 789–797. doi: 10.1038/mp.2010.8

Kotov, R., Chang, S.-W., Fochtmann, L. J., Motjtabai, R., Carlson, G. A., Sedler, M. J., et al. (2011). Schizophrenia in the internalizing-externalizing framework: a third dimension? *Schizophr. Bull.* 37, 1168–1178. doi: 10.1093/schbul/sbq024

Kraepelin, E., and Diefendorf, A. R. (1907). *Clinical Psychiatry: A Textbook for Students and Physicians*. 7th Edn. London: MacMillan Company.

Lazarewicz, M. T., Ehrlichman, R. S., Maxwell, C. R., Gandal, M. J., Finkel, L. H., and Siegel, S. J. (2010). Ketamine modulates theta and gamma oscillations. *J. Cogn. Neurosci.* 22, 1452–1464. doi: 10.1162/jocn.2009.21305

Lee, S. H., Ripke, S., Neale, B. M., Farame, S. V., Purcell, S. M., Perlis, R. H., et al. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* 45, 984–994. doi: 10.1038/ng.2711

Ma, L., Skoblenick, K., Seams, J. K., and Everling, S. (2015). Ketamine-induced changes in the signal and noise of rule representation in working memory by lateral prefrontal neurons. *J. Neurosci.* 35, 11612–11622. doi: 10.1523/JNEUROSCI.1839-15.2015

MacDonald, A. W., Ill. (2013). “What kind of a thing is schizophrenia? Specific causation and general failure modes,” in *Schizophrenia: Evolution and Synthesis*, eds S. M. Silverstein, B. Moghaddam, and T. Wykes (Cambridge, MA: MIT Press), 25–48.

Markon, K. E., Chmielewski, M., and Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. *Psychol. Bull.* 137, 856–879. doi: 10.1037/a0023678

Neyminton, S. A., Lazarewicz, M. T., Sheft, M., Contreras, D., Finkel, L. H., and Lytton, W. W. (2011). Ketamine disrupts theta modulation of y in a computer model of hippocampus. *J. Neurosci.* 31, 11733–11743. doi: 10.1523/JNEUROSCI.0501-11.2011

Rausan, M., and Hoyland, A. (2004). *System Reliability Theory: Models, Statistical Methods and Applications*. 2nd Edn. Hoboken: John Wiley and Sons.

Rolls, E. T., Loh, M., Deco, G., and Winterer, G. (2008). Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat. Rev. Neurosci.* 9, 696–709. doi: 10.1038/nrn2462

Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinsen, R. K., et al. (2010). Developing constructs for psychopathology research: research domain criteria. *J. Abnorm. Psychol.* 119, 631–639. doi: 10.1037/a0020909

Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427. doi: 10.1038/nature13995

Spitzer, R. L., Endicott, J., and Robins, E. (1975). Research diagnostic criteria. *Arch. Gen. Psychiatry* 32, 777–781. doi: 10.1001/archpsyc.1972.01750190059011

Spitzer, R. L., Endicott, J., and Robins, E. (1975). Research diagnostic criteria. *Arch. Gen. Psychiatry* 32, 777–781. doi: 10.1001/archpsyc.1972.01750190059011

System Reliability Theory: Models, Statistical Methods and Applications. 2nd Edn. Hoboken: John Wiley and Sons.

FUNDING

This work was supported in part by NIMH grant R01MH084861 (PI: AM3) and NSF career award 1135581 (PI: TN).

Copyright © 2016 MacDonald, Zick, Chafee and Netoff. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction which does not comply with these terms.