The Objectivization of Medical Diagnosis: The Use of Machine Learning Techniques in Magnetic Resonance Imaging

---

A thesis submitted in partial fulfillment of the requirements for Honors in the Science and Society concentration of Brown University.

---

by Richard Fadok

Advisor and First Reader: Catherine Bliss, PhD

Second Reader: Sherine Hamdy, PhD

Brown University

April 29th, 2011

Table of Contents

Acknowledgments

The inception of this thesis, nevermind its completion, would not have been possible without the support of Professor Catherine Bliss. From the first paper I wrote for her eye-opening class my junior year to this present day, she has always made feel that my ideas were unique, insightful, and legitimate. At Brown, she has been my greatest mentor, urging me to corral my spatially and temporally scattered musings into well-formed academic contributions. I attribute to her my interest in STS and my decision to study biomedicine and society next year at LSE, and for those reasons I want to thank her first and foremost.

I also want to give special thanks to other members of the STS community. Professor Sherine Hamdy was kind enough to agree to read my thesis at a very late date. She also provided invaluable feedback at my presentation that will guide my future inquiries into this field. In addition, I want to thank Professor Anne Fausto-Sterling for allowing me to declare Science and Society as a second concentration fairly late in the game and for allowing me freedom in constructing the concentration requirements. Her faith in me did not go unnoticed.

I would like to thank several other amazing professors who have influenced how I think critically: Vanessa Ryan, Marinos Pourgouris, Barbara Herrnstein-Smith, and Wendy Chun. In particular, I want to thank Professor Pourgouris and Professor Herrnstein-Smith for teaching me that our most playful thoughts can be our most beautiful.

There are countless friends and family members that I would like to thank. In short though, I want to thank my best friends, Alex Wong and Jonnie Lappen. They have been some of my greatest teachers. I also want to thank my mom and dad. They have never wanted anything other than for me to pursue what makes me happy, and I can't thank them enough for this.

Lastly, I want to express my thanks and love for Scott Gallagher. More than anyone, he has made me believe in myself: in the value of my ideas and especially in my ability to complete two theses this year. He has calmed me down when I have panicked at midnight, helped me collect my thoughts several hours later, and laughed with me after deadlines when we look back and realize how unreasonable I was being. I want to dedicate this thesis to him.

Abstract

A growing body of magnetic resonance imaging research focuses on the development of machine learning techniques for the diagnosis of neurological and psychiatric disorders. Workers in this field trumpet machine learning techniques, a set of algorithms that enable the automated classification of data into classes, as the means to make diagnostics an objective science. By delegating the task of diagnosis – the classification of persons as healthy or diseased – to the automated machine, it is believed that the dubious traces of human subjectivity can be eradicated from diagnosis, thereby ensuring the purity of the disease's truth. In this thesis, I examine the machine learning technologies and conduct a discourse analysis on scientific literature to interrogate the beliefs surrounding these techniques.

My work reveals that, despite claims of objectivity, the machine learning algorithm is never completely separated from the human: Machines are complex fusions of humans and nonhumans that work together to generate the truth of a disorder. This hybridity is obscured, however, by the researchers, whose discourse produces an ontological difference and hierarchy in constitution and functioning between humans and machines. Distinguishing the machine from the human, and translating the human production of knowledge into a machine task, is a rhetorical practice that authorizes the truth of disease. Locating the traces of the human and destabilizing the purity of the machine as object is a necessary first step to subjecting these discourses of power to critique.

For Scott

"So long as humanism is constructed through contrast with the object that has been abandoned to epistemology, neither the human nor the nonhuman can be understood."

(Bruno Latour, 1993)

Introduction

Objectivity holds a deep fascination for the practice of medicine, with clinicians and researchers alike holding objectivity to be the precursor to uncovering the truth of an illness. Diagnosticians of the brain-self – neurologists and psychiatrists – especially strive to produce objective knowledge. Neuroimages, for example those obtained by magnetic resonance imaging (MRI), have been heralded as the technical means to obtain this lofty desideratum, for their mimetic visual relationship with unaided sight produces the feeling of apprehending a reality truly independent of the viewing subject.

In spite of this turn to neuroimaging, the fear of an intervening subject still arrests the profession. Behind every diagnosis, even one based off an examination of a neuroimage, lies a human diagnostician, out of whom leak whims, ideologies, prejudices, and desires that threaten to contaminate and stain the purity of the objective representation. In an effort to overcome this last obstacle standing between them and the moment of truth, researchers in psychiatry and neurology have been investigating a class of novel statistical techniques, borrowed from computer science and widely referred to as machine learning, that enables the automatic classification of data. By applying these techniques to MRI, they are developing the means for the automated detection of brain-based illnesses. This way it is believed that the last remnant of subjectivity can be extracted from neurology and psychiatry.

Machines, however, do not spontaneously spring into existence. They are artifacts of human construction, MRI machine learning algorithms being no different. From their design to their execution, these algorithms are messy imbroglios of nonhumans and humans, of cogs and cognition. Nevertheless, these hybrid creatures are invariably reduced to pure, mechanical objects. By "black box"-ing these machines (Latour 1987), the human figure could be erased and the knowledge objectivized, a process that leaves it bodiless, elusive, and inscrutable. To unearth the human again, an archaeology of the human at the site of the machine is necessary.

My research attempts to answer two questions. First, where is the human in the machine learning algorithm? Second, why has the figure of human been obscured? In this thesis, I seek to un-“black box” these machines to ground the knowledge they produce in a human body. To do this, I analyze the machine to unearth and illuminate the human presence. I also present discourse analysis on the representations used by scientists in the machine learning literature to investigate why objectivity is perceived to be desirable. By tracing the diagnosis, I want to make it responsible for its effects and amenable to discussion and debate.

Machine Learning: A Primer

Machine learning, in which new data is analyzed by an algorithm built up from old data, is no newcomer to the realm of biomedicine. For years, machine learning techniques have been incorporated into bioinformatics research, particularly for their ability to identify and classify genomic and proteomic biomarkers in an array of diseases. Interest in these seductive automatons has recently spread to neuroimaging research, where they have been modified and translated for application to brain-computer interfaces, brain state prediction, identification of distributed network activity in the brain, and automation of brain segmentation. Interest in the use of machine learning for disease diagnosis and prognosis, however, is a recent phenomenon.

Diagnosis by neuroimaging occurs in only a handful of neurological and psychiatric disorders. In neurology, a diagnosis is typically reached through a combination of patient history, neuropsychological examination, and visual inspection of neuroimaging results, usually computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) scans. In psychiatry, neuroimaging is almost never used for diagnosis, which is instead attained through a clinical assessment of behavioral symptoms and patient self-report. The novelty of machine learning techniques for neurological and psychiatric diagnosis lies in the shifting locale of the truth of disease, from "subjective," symptom-based medicine to "objective," biomarker identification, from presentation to essence, and from clinical observation to mechanical automation.

The use of machine learning for disease diagnosis and prognosis, the focus of my research, is prevalent in all modalities of neuroimaging technology, including near-infrared spectroscopy, electroencephalography, magnetoencephalography, magnetic resonance spectroscopy, microscopy, CT, and PET. Fascination with machine learning for the diagnosis and prognosis of neurological and psychiatric diseases has manifested most intensely, however, in the field of MRI. Possibly due to its status as a salient cultural symbol (Joyce 2008), its presence in most biomedical research and treatment facilities, and the versatility with which it can be manipulated with to produce different types of information, MRI is a dominant player in machine learning-based diagnosis. Although still relatively young, machine learning has already been applied to the diagnosis of numerous neurological disorders: Alzheimer's disease, mild cognitive impairment, and frontotemporal lobe degeneration (Anderson et al. 2010; Avants et al. 2010; Chou et al. 2009; Costafreda et al. 2011; Cuingnet et al. 2010; Davatzikos et al. 2008; Fan et al. 2008; Ferranini et al. 2008; Filopoyv et al. 2011; Gerardin et al. 2009; Hinrichs et al. 2010, 2009; Hua et al. 2010, 2009; Leung et al. 2010; Li et al. 2007; Magnin et al. 2011; Isra et al. 2009; Morra et al. 2009, Morra et al. 2008b, 2008a; Plant et al. 2010; Stonington et al. 2010; Toews et al. 2010; Vemuri et al. 2011, 2007; Wee et al. 2011; Westman et al. 2011; Young et al. 2009; Zhang et al. 2011), dementia (Burge et al. 2009; Chen et al. 2010), brain tumors (Blanchet et al. 2011; Emblem et al. 2011, 2010; Georgiadis et al. 2011, 2009; Hu et al. 2011; Juntu et al. 2010; Kassner et al. 2010; Zacharaki et al. 2009; Zollner et al. 2010), epilepsy (Duchesne et al. 2006; Kassner et al. 2010; Kodipaka et al. 2007; Raj et al. 2010), movement disorders (Focke et al. 2010); traumatic brain injury (Perlbarg et al. 2009), multiple sclerosis (Lin et al. 2006), aphasia (Wilson et al. 2009), Huntington's Disease (Rizk-Jackson et al. 2010), and autism (Ecker et al. 2010; Jiao et al. 2010). Machine learning techniques have also been advanced within psychiatry and have been applied to schizophrenia (Anderson et al. 2009; Ardekani et al. 2011; Guo et al. 2008; Kasparek et al. 2011; Shen et al. 2010; Yoon et al. 2007), depression (Craddock et al. 2009; Gong et al. 2010; Nouretdinov et al. 2010; Qing et al. 2010), and obsessive-compulsive disorder (Soriano-Mas et al. 2007). This section will trace the scientific emergence of these diagnoses and prognostic prediction machines, beginning with hydrogen atoms and ending with the label of "diseased" or "healthy."

To begin with, subjects are placed in a scanner, which artificially induces an incredibly powerful magnetic field (on the order of sixty thousand times the strength of the Earth's magnetic field) that is used to align the various spins (spin is a fundamental physical property like mass that describes the angular momentum of a particle) of hydrogen atoms in the brain. Afterwards, a pulse of radio frequency energy is applied briefly to perturb these spins out of alignment. Due to the intrinsic propensity for these spins to realign back with the magnetic field, at different points in time depending on their local molecular environment (e.g., attachment to proteins, fats, or carbohydrates; membership in free-floating water molecules; temperature; etc), MRI scientists can infer some of the spatial and chemical environment of the brain by using receivers that detect the signal, an electrical current, generated by this realignment. Numerous components of this procedure can be manipulated by MRI technicians, including the strength of the scanner; the atom under investigation (in this example, hydrogen was used); the frequency, strength, and duration of the pulse applied; and the application of various other sequences and components that are used to identify the three-dimensional characteristics of the signal.

By modifying the scan sequences, MRI can be adapted to a variety of purposes. Researchers and clinicians interested in analyzing the structure of the brain can employ several different scan sequences to highlight or downplay various components of the local environment. Functional MRI utilizes a different scan sequence than structural MRI to capitalize on the differences between the magnetic properties of oxygenated and deoxygenated blood. In doing so, activity in the brain related to cognitive, emotional, and autonomic functions can be characterized. A third variant of scan sequence that figures heavily in the machine learning literature was diffusion tensor imaging (DTI). In tissues that exhibit structures that facilitate rapid water diffusion along one axis better than others (i.e., axons, the long cellular projections of neurons), directionality of diffusion can be inferred using the principles of magnetic resonance discussed above. In this manner, the tracts of white matter axons can be derived and connected to create information related to the structure of the brain's neural connections.

After data acquisition, the raw measurements are digitized and subjected to a series of processing manipulations, which allow for the researcher to decide what does and what does not appear in the data. Although some of these "pre-processing" steps are automated, a researcher must ultimately decide which algorithms to choose; furthermore, many of them are not. Data correction takes place first to account for unwanted head motion, magnetic field inhomogeneities, random "noise" (cellular or digital activity that is not deemed largely relevant), distortions of the pulse, blurring due to air pockets (particularly around the sinuses, which influences visualization of some of the brain regions near the temporal bones), and other physiological and technical artifacts. Sophisticated algorithms have been derived to account for these artifacts, but, importantly, not all of them are removed (or even understood); and all of them involve interpretation, which is never innocent or uninformed. The remaining steps vary from study to study, but typically the data from an individual are aligned with a template provided. These templates represent idealized brains in various forms; some are averages while others are from a single individual. In both cases, the spatial coordinates of individual brain space are transformed to match an idealized template brain so that, despite individuality variability in brain structures, mass comparisons can be made between populations.

Typically, the voxels (*vo*lumetric pi*xels*) would then compared for statistical differences between groups based on the prior hypotheses of the researchers. Then, an image would be produced based on these statistically significant differences and read by a diagnostician for information regarding disease. Machine learning takes a different, automated approach, one that does not result in an image but rather a binarized decision.

Because the amount of data is too massive to analyze at once, a set number of features (distinct subsets of data that could be artifactual or relate to a physiological process, brain structure, or temporal link between activities in the brain) to work with is desired. Numerous algorithms exist to reduce the dimensionality of the data and select a group of features for further analysis, including principle component analysis, independent component analysis, and common spatial pattern; other times, the features of interest are identified by the researcher. After this dual process of reduction and selection, a pre-selected classification algorithm is deployed that divides the data based on the features into distinct classes that exhibit statistical differences. The algorithms that one can find in the neuroimaging literature are linear discriminant analysis, Fisher's discriminant analysis, support vector machines, K-nearest neighbors, fuzzy logic, and random forests. A thorough discussion of the precise functioning of these algorithms is beyond the scope of this analysis. It will suffice to say that, by using a set of "training" data that is fed into the machine, these algorithms mutate to learn what separates the data most clearly (i.e., with minimal error). The models are then analyzed by third type of algorithm that tinkers with the algorithms' generalizability and reliability until they can be optimally be used to predict the output of new data, which, without knowing its class prior to input, can be classified and, in the case of disease state, diagnosed (Lemm et al. 2009; Pereira et al. 2009). So, in concrete imagery, a group of "training" patients are initially scanned. From these patients, an algorithm is developed that can then be used to classify new patients as diseased or healthy.

Data and Methods

In this thesis, I first document and then dissect the ways in which machine learning is represented as an objective means to reach the truth of medical diagnoses using MRI. I employ discourse analysis to attend to the rhetoric that circulates within and around machine learning discussion. Because the technique is so new, there exist few appearances of machine learning in the popular media. Therefore, in order to demystify the technique and complicate the efforts at "black box"-ing machine learning (Latour 1987), I probe scientific literature for information regarding the precise nature of the methods. The extrapolations and observations scientists drew in their conclusions are then investigated. With this dual approach, the process of objectivization –the construction and crystallization of "objective" knowledge – can be made visible.

I selected the journals by referencing the Journal Citation Reports, a database compiling the leading publications in various academic disciplines and ranking them by their impact factor, a measure of the number of citations per article in a given publication. Impact factor did not figure into the choice of journals, however, as all listed under the heading of "Neuroimaging" were targeted for completeness. Thirteen in total were selected, and after removing three without access from Brown University's Library Services (*Neuroimaging Clinics of North America*, *Minimally Invasive Neurosurgery*, and *Journal of Neuroradiology*) and one that was exclusively dedicated to a non-MRI neuroimaging technique (*Clinical Electroencephalography in Neuroscience*), nine remained. These included *Human Brain Mapping*, *Neuroimage*, *Psychiatry Research – Neuroimaging*, *American Journal of Neuroradiology*, *Neuroradiology*, *Journal of Neuroimaging*, *Stereotactic and Functional Neurosurgery*, *Brain Imaging and Behavior*, and *Clinical Neurophysiology*. In addition to these nine journals, an additional four were added that were dedicated to MRI research in imaging technique and clinical translation. These included *Journal of Magnetic Resonance Imaging*, *Magnetic Resonance in Medicine*, *Medical Image Analysis*, and *Magnetic Resonance Imaging*.

I picked relevant articles from these journals using the search terms "machine learning" and "support vector" (the most common machine learning technique) within the time frame of January 1st, 2005 and February 22nd, 2011. After this initial search, I skimmed through and screen in 408 articles if they included the use of machine learning for the automated diagnosis of any neurological or psychiatric disorder in humans using MRI. Out of 408, 69 articles passed this initial screen. I then scrutinized these articles for a detailed understanding of the technique and for the rhetoric used to describe machine learning, with an emphasis on surrounding claims of objectivity.

Speaking for Machines

Narratives of objectivity dominate the discourse on machine learning found in scientific publications. In the literature, machines are depicted as totally objective, and humans subjective. Assuming a completely mechanical process preserved from the spoiling influences of subjectivity – its conceptual antithesis and dark obverse – objectivity is posited as the ideal towards which diagnostics should strive. In almost all of the articles I analyzed, researchers explicitly or implicitly sought to render their methods and depict their results as objective. This process of objectivization frequently entailed the construction of a clean and hierarchical dichotomy between the human and the machine. When a given procedure, from tissue segmentation to diagnosis, had been completely mechanized, its presence in the text was accompanied by a number of descriptive qualifiers that elevated the position of the machine to one of superiority compared to the human. This next section will document the rhetoric that scientists use in describing machine learning as objective.

Difference

A foundational principle for machine learning researchers is that the human and the machine are distinct and self-contained entities. Rarely articulated but always tacitly present, the human/machine dichotomy implies a clear ontological difference between machines and humans. Machine learning scientists definitively understand a particular methodological step to be either machine-driven or user-dependent. Although there exist methodologies that are a mixture of human and machine components, the individual components are always discretely represented as human *or* machine. The totally objective process is pure machine. As one research group starts their paper, "The general goal of our work is to develop a classification method that is independent of any trained user interaction" (Anderson et al. 2009: 2510).

The independence and purification of the machine from the human is construed to be both possible and desirable. This same team later wrote, "Our method operates using all the independent components within a subject, so no human interpretation is required to achieve classification of the data" (Anderson et al. 2009:2517). Another group similarly distinguishes between raters (the humans) and the automated process (the machine), "We exploit the potential of features extracted automatically from the images in order to avoid descriptive criteria, which are rater dependent" (Zacharaki et al. 2009:1610). Conceptually, the scientists perceive automated machines to be independent of humans, and they seek to actualize this perception in their practices.

Hierarchy

In addition to a division, there exists also a hierarchy between the two members of the human/machine dichotomy in the literature. Machines are often portrayed as superior, evidenced in one author's conclusion, "The main limitation of this approach is its dependence on a human analysis of the K-means results. Indeed, an operator must select the classes representing the tumor core" (Blanchet et al. 2011:72). The author believed the major flaw of the study was its reliance on "human analysis," suggesting that machines are an improved means to conduct the experiment, with human intervention being anathema. Objectivity is associated with the absence of any manual interaction (Wu et al. 2006). Empirical results of several studies also support this hierarchy, with the following representing an average response, "The classification accuracy of the SVM [support vector machine], which was 93%, was better than the radiologist classification accuracy" (Juntu et al. 2010:680). Human interpretation is seen as inferior, a fact that many scientists believe to be validated by the results obtained in their studies.

Variability

The most common critique of human methods was that they are subject to error, bias, and variability. The machine, however, is celebrated as an iterative and indefatigable method, producing perfect, reproducible, and untainted knowledge. For one group, objectivity is a function of the iterability of a mechanized process, "The objectivity of texture analysis depends on the assumption that images are acquired, processed, and analyzed under identical conditions" (Kassner et al. 2010:812). For another, subjectivity is a major source of inaccuracy and error, "The subjective nature of many of the decisions related with the process of brain tumor characterization has led clinicians to continuously seek for greater accuracy in the characterization of brain tumor tissues" (Georgiadis et al. 2009:120). The productive capacity of humans is limited by their tendency towards change and variability; the mechanical constancy of the machine is the solution to these perceived problems. Kassner et al. (2010) sums up best the sentiment shared by many machine learning researchers:

> Conventionally, radiologists produce diagnoses on the basis of a combination of their training, experience, and individual judgment. Radiologists perceive and recognize image patterns and associate or infer a diagnosis consistent with those patterns. It follows that there will be an inevitable degree of variability in image interpretation as long as it relies primarily on human visual perception. Tools for automated pattern recognition and image analysis can provide objective information to support clinical decision-making and may serve to reduce this variability (p. 809).

Here, human perception and judgment are represented as, by their very nature, susceptible to error and variability that the machine is impervious to. Because of their different essence, machines generate diagnoses that are superior to those the human doctor is capable of.

Invisibility

More than just their perceptual variability, humans are also cast in the literature as perceptually limited compared to machines. Many patterns that are present in the data are "imperceptible to the human visual system" (Kassner et al. 2010:809). On the contrary, machines "might have the potential to discriminate between some cases of malignancy that humans cannot recognize easily by visual inspect" (Juntu et al. 2010:681). Because of the intrinsic restrictions on their visual capabilities, humans are subject to a higher risk of misinterpretation and error when diagnosing. But humans are not just limited in their ability to view the visible; machines have the ability to see beyond, to apprehend the unseen, "spectral" (Kassner et al. 2010), and invisible form of the data. Even when no symptoms are even present, the machine can perceive illness, as Qing et al. (2010) explains:

> The clinician must choose the answer to each question by interviewing the patient and by observing the patient's symptoms. As a result, such a subjective diagnosis procedure is sensitive to various unpredictable disturbances from outside, especially in the early stage when the signs and symptoms are recessive (p. 1067).

Whether symptoms are recessive, or present but, due to humans' limited seeing apparatus, imperceptible, machines are more visually adept.

Speed

Another consistently described advantage of machines was their enhanced speed. For machine learning researchers, controlling time through the increased processing velocity of machines was essential to their program. Mastering time meant mastering subject flow, as one researcher notes, "These techniques offer promise from improved clinical workflow, including clinical research studies such as longitudinal monitoring of the evolution and treatment of degenerative and inflammatory diseases" (Mortamet et al. 2009:365). Improved workflow permits researchers to expand the number and types of studies that they conduct. Automation "accelerate[s] epidemiological studies and clinical trials" (Morra et al. 2008:60) and eliminates prohibitions on knowledge as a result of labor-intensive methods (Yoon et al.

2007). For clinicians too, time was a major issue because increased flow meant higher revenue (Kassner et al. 2010). Machines sped up the diagnostic process, which Zacharaki et al. (2009) remarks on, “Automated tools, if proven accurate, can ultimately be applied to…(iii) expedite or anticipate the diagnosis (histologic examination is usually time consuming” (p. 1609). The application of machine learning to diagnosis is the ultimate means to accelerate and control flow, whether patient or participant, because “once the system has been designed, no iterative processes are required in order to classify a new case. Thus, classification of new cases is instantaneous” (Georgiadis et al. 2009:129). By approaching and often times reaching the speed of instantaneity, machines control time, making humans appear sluggish and inefficient.

Progress

A grander meta-narrative of technological and intellectual progress seems to characterize much of the discussion on machine learning. Three indicative examples demonstrate this. First, Qing et al. (2010) opposes the automated method to tradition, “Compared with traditional diagnosis methods, such as HAMD, the model is objective and easy to use” (p. 1073). Second, Blanchet and colleagues (2011) differentiates automation from *classic* methods, “The case would have been identified by classic inspection of the data” (p. 71). Third, Wu et al. (2006) writes, “Most previous methods were labor-intensive, subjective, and provided little if any anatomic localization” (p. 141). In these texts, it is insinuated that objectivity is modern, progressive, and enlightened; in contrast, subjectivity is primitive, antiquated, and inumbrated.

Agency and Critique

In the majority of the articles I examined, machines were attributed with agency.

Machines could learn, respond to training, promise, anticipate, distinguish, support, assist, confirm, deny, analyze data, and create diagnoses. Like humans, machines were endowed with the ability to act, to perform, and to know; they were participants in the same scientific process that humans were. Machines produced knowledge. Unlike the knowledge issued from humans, however, the knowledge generated by machines was objective. In the results sections, knowledge is not produced by the scientists involved in the study but rather by the machine. It is during the discussion that humans resume their role as knowledge-producers and express their thoughts. Occasionally, machine-generated knowledge is criticized, but, as discussed above, this is proportional to their dependence on humans; completely objective, human-independent knowledge was immune from criticism. Thus, while both machines and humans are agents in the scientific process, machines resist and deflect critical discussion.

Quantitativeness

Another distinctive characteristic of machines, as presented in their literature, was their inherent propensity for objective, quantitative measurement. Humans were, by definition, qualitative. Description was subject to variability and imprecision while numbers were universal, invariable, and final. Machine learning provided researchers the ability "to characterize the disease process in terms of objective, quantitative measurements of brain function and anatomy" (Avants et al. 2010:1005). When applied to diagnosis, "an accurate prediction model would provide a quantifiable and objective framework to assist the currently symptom-driven subjective treatment decision-making process for schizophrenia" (Guo et al. 2008:1106). Symptom-based diagnosis was inaccurate, qualitative, and obstructive to objective disease assessment.

Cyborg Diagnostics

Scientific representations of machine learning depict machines as objective technologies that can reveal the truth of an illness by completely removing the human from the procedure. Because humans are variable, prone to error, and biased, they are liable to infuse their own whims and desires into a diagnosis. Eliminating subjective, human influences through automation is perceived to be the path by which the disease can be revealed as it really exists, independent of human observers.

Daston and Galison (2007) have extensively traced a genealogy of objectivity and demonstrated that it was not always, as it has come to be thought of in recent years, an intrinsic property of any representation but was, and remains, a regulative technique of the self – a scientific virtue – that disciplines and conditions the scientific practitioner. As a normative code of both being and knowing, objectivity delimits not only what can be known about the world but also how it can be known.

Today's form of objectivity is not the only scientific virtue, however; other values have prevailed in previous years. Prior to the eighteenth century, the prevailing epistemic virtue was "truth-to-nature." Adherence to this stance implied that scientists concerned themselves with seeing beyond the sometimes-monstrous particularities and oddities through which nature presented itself. In response to this Manichean nature that endeavored to befuddle the rational observer, a specific type of scientist was cultivated, one who possessed the experience, wisdom, and reason that would aid him in extracting the pure form of nature from its variegated manifestations. The implementation of this technique therefore "required that they actively select, sift, and synthesize the sensations that flooded the too-receptive mind" (Daston and Galison 2007:203). Truth-to-nature necessitated the mediation of the scientist, for his very activity was crucial to the rigorous discernment of truth that was valued in this allegiance to "truth-to-nature."

In contrast to this regime stands mechanical objectivity. In the eighteenth century, Immanuel Kant ushered in a new mode of self-training with his philosophy of a unified and active sense of self. After Kant's reintroduction of "objectivity" and "subjectivity" into the philosophical register, this new epistemic virtue arose to oppose that of "truth-to-nature." Moving away from theories of both the rational soul and the associationist mind, post-Kantian subjectivity "presumes an individualized, unified self organized around the will" (Daston and Galison 2007:33). Whereas the "too-receptive" pre-Kantian self demanded the active force of reason to shelter it from the stupefying deluge of incoming stimuli, "those who deployed post-Kantian notions of objectivity and subjectivity had discovered a new kind of epistemological malady and, consequently, a new remedy for it" (Daston and Galison 2007:33). This malady was not the self's tendency to be overwhelmed but the subject's inherent propensity for spilling out of itself – for impressing itself upon the objective world.

As a result of this reformulation of the self, the era of mechanical objectivity saw a marked self-abnegation on the part of the scientist. If the evils that the true-to-nature scientist fought off consisted of the impressionability of the self, the new vices "to be resisted were the temptations of aesthetics, the lure of seductive theories, the desire to schematize, beautify, simply – in short, the very ideals that had guided the creation of true-to-nature images" (Daston and Galison 2007:120). Because the very willfulness of the subject was perceived to be an obstacle to the pursuit of truth, it had to be quarantined and eliminated. Thus, the scientist who sought after objectivity barred his own subjectivity from leaving any mark on the representations he generated: "To be objective is to aspire to knowledge that bears no trace of the knower" (Daston and Galison 2007:17).

Mechanical reproduction provided the vehicle to reach this goal. By delegating the task of selfless representation to the indefatigable machine, objectivity could be attained. The irregularity and multiplicity of nature could now be faithfully and objectively rendered without anxiety about infection by the subject. With the advent of automaticity came an entirely different scientist from before. As Daston and Galison (2007) remark, "What characterized the creation of late nineteenth-century pictorial objectivism was self-surveillance, a form of *self*-control at once ethical and scientific…scientists came to see mechanical registration as a means of reining in their *own* temptation" (p. 174). However, as with truth-to-nature before it, it was not long before the tenets of mechanical objectivity began to come under attack from its successor, trained judgment.

As some scientists began to realize and attempt to address the insufficiencies of a strictly mechanical paradigm of objectivity, the virtue of trained judgment was born. The critique of mechanical objectivity that proponents of trained judgment mounted was a direct confrontation with one of objectivity's central ideals, the total absence of any interpretation. Especially in the medical profession, the problem was posed: "If one is committed…to the mechanical registration of images of individuals, then how can one distinguish between variations within the bounds of the 'normal' and variations that transgress normalcy and enter the territory of the pathological?" (Daston and Galison 2007:309). Recognizing the limitations of mechanical objectivity, trained judgment stressed "interpolation, highlight, abstraction – all were subtle interventions needed to elicit meaning from the object or process, and to convey that meaning – to teach expertise – through the representation"

(Daston and Galison 2007:348). The self that was denied under the rule of objectivity returned full force as a necessary prerequisite to the discovery of truth and knowledge.

From truth-to-nature to trained judgment, the scientific self has undergone a great deal of transformation. The emergence of each epistemic virtue has brought with it new normative modes of being and knowing in order to correct the errors of its predecessors. But Daston and Galison (2007) emphasize:

> This history is one of innovation and proliferation rather than monarchic succession…Each successive stage presupposes and builds upon, as well as reacts to, the earlier ones…This is not some neat Hegelian arithmetic of synthesis, but a far messier situation in which all the elements continue to play and in interaction with one another…In contrast to the static tableaux of paradigms and epistemes, this is a history of dynamic fields, in which newly introduced bodies reconfigure and reshape those already present, and vice versa (p. 18).

Within the field of neuroimaging, STS scholars have demonstrated that virtues of both trained judgment (Prasad 2006) and mechanical objectivity (Beaulieu 2001) exist simultaneously as disciplining scientific norms.

Prasad's (2006) ethnographical study of how neuroimages are used in reaching a diagnosis provides evidence that one of the prevailing epistemic virtues is trained judgment. Prasad reveals that, in addition to the objective neuroimage, several other knowledge-sources (including the physician's history, the patient's history, and a host of other physiological tests) are drawn from to determine the truth of a patient's condition. As a whole, the diagnostic process is, therefore, an exercise in trained judgment, for the emphasis is on the physician's "expert training of the eye" (Daston and Galison 2007:331). Thus, neuroimages do not provide the only truth, and the knowledge about a patient's condition is not entirely objective: The particular position of the physician is fused with the information gleaned from the neuroimage. Prasad (2005) writes:

> Practioners seek to locate pathology through a 'differential analysis' of these diverse sets of MR images. The differential analysis/viewing is possible because of a dynamic interaction between the scientist/radiologist and the image data that would not be possible without the help of a computer. The construction of new images using MRI is intrinsic to the radiological interpretative process. Closure on pathology is achieved, however, not only through differential analysis of the images but also through cross-referencing different 'inscriptions'—images, diagnostic data, and so on, which together constitute the radiological gaze and function to detect and fix pathology (p. 292).

At least in the radiological practice, neuroimages are not divorced from a human perspective. They are not the only source from which a truth is vocalized. Rather, numerous channels of truth weave and intersect in this "differential analysis." At the end of all of this, it is up to the doctor to negotiate which are more true and accurate than the others. At its core then, this is a process of negotiation, in which the objective reality of a depiction in a neuroimage might be supplemented, *or even supplanted,* by other forms of knowing and seeing. Prasad (2006) argues that this is evidence for a new form of visuality that he calls "cyborg visuality," based off an understanding of a fusion between man and machine provided by Haraway (1991).

While Prasad's (2006) study certainly proves the existence of the virtue of trained judgment, Beaulieu's (2001) articulation of "digital objectivity" indicates that mechnical objectivity has not been entirely ousted as a scientific virtue in the field of neuroimaging, and has even refashioned itself in light of other virtues. For Beaulieu, the digitization of nature into the neuroimage does indeed recast mechanical objectivity in light of trained judgment: The introduction of reliability estimates during image processing steps channels the worries of trained judgment into a mechanical process. Beaulieu (2001) writes, "Whereas the traditional atlas is heuristic, and the observer (surgeon or anatomist) must make a judgement [sic] call in applying her knowledge, in the new digital atlases this process is automated…these atlases also offer the possibility of quantifying the experimental error" (p. 663-664). Importantly, the goal of mechanical objectivity is still reached after this "orthogonal innovation" since "the observer hardly appears at all in digitalized work based on a virtual object…and when she does, it is then only to test the pipelines, not the accuracy of the transformations – for which there are other automated testing methods" (Beaulieu 2001:664). Thus, mechanical objectivity has been transfigured into digital objectivity and co-exists in dynamic tension with trained judgment vis-à-vis cyborg visuality.

Previous STS scholarship has done tremendous work in uncovering the nature of objectivity in MRI; however, the introduction of machine learning techniques in diagnostic MRI requires updated analysis. Prasad's analysis privileges the eye: The vision of the radiologist is what interacts with the objective image to create a new ethic of knowledge. Within machine learning literature, however, there is no viewing subject. The machine learning algorithm generates a "yes" or "no" answer in response to the question of whether disease is present. Because no image is produced, no human views it; the machine subsumes this entire process. As a result, machine learning research would appear to be more in line with Beaulieu's analysis; mechanical objectivity increasingly dominates in machine learning-based diagnosis as more and more processes become automated.

While Beaulieu's analysis holds to some extent in terms of her analysis of the machine's output, the ideated role of the machine in the diagnosis suggests that this is not the full story. As I found in my examination of the literature, researchers do not propose that machine learning diagnostics should replace the role of the doctor. Instead, they argue, for the most part, that the result generated from the machine should add to, assist, enhance, complement, aid, support, guide, and supplement the human diagnosis. Vemuri et al. (2008) conclude their analysis:

> Sound clinical validation is a necessary first step. Subsequent studies will evaluate clinical utility where the clinical/psychometric evaluation does not provide an obvious answer. This algorithm could be used as a stand alone application or more likely, in conjunction with evaluation methods employed in standard clinical practice (p. 1194).

"Clinical utility" will depend not on one method alone but with a combination of all clinical methods. Guo et al. (2008) also adopt this standpoint, "This pragmatic shortcoming suggests the utility of developing a statistical framework to predict treatment-related brain alterations, which could combine with baseline scans and patient history to inform clinical decision-making" (p. 1093). Although several studies do indicate that the machine diagnosis will replace the human diagnosis, the majority believes that machine learning techniques will only add a second (or third, fourth, fifth, etc) opinion to the host of evaluative techniques the doctor already implements.

Importantly, the machine does not *supplant* the human. In other words, the application of machine learning does not eliminate the position of the doctor, as fears of technological domination would anticipate. Curiously, despite criticisms of the subjective flaws related to human analysis, a complex relationship between human and machine is advocated, one in which they collectively endeavor to uncover the truth of the diagnosis. Thus, I propose that, instead of viewing them as instances of pure digital objectivity, or "cyborg visuality," we should consider diagnosis via machine learning as *cyborg diagnostics* – complex hybrids of man and machine acting to produce knowledge about the truth of a patient's diagnosis. Moreover, neither man nor machine dominates the action (Latour 1999), but both work together to achieve a new diagnosis that was not possible without the agency of each.

Another crucial problem limits the utility of Daston and Galison's, Beaulieu's, and Prasad's analyses of objectivity: the absence of a recognition of the human constructedness of the machine. Although the artificiality of the neuroimage has been greatly studied (Joyce 2008; Dumit 2004), the machine learning algorithm has received no attention. For objectivity theorists, objectivity appears to be some perfect, harmonious state that can be realized with varying degrees of success. Mechanical objectivity is a pure ideal, "never perfectly attained but consequential all the way down to the finest moves of the scientist's pencil and the lithographer's limestone" (Daston and Galison 2007:143). Moreover, Beaulieu (2001) claims that "the observer hardly appears at all in digitalized work" (p. 664) and that this might signify "the extent to which the pole of the 'subject' has been diminished in this context" (p. 664). The problem inherent in each of these interpretations is the belief in the purity of the machine; once mechanized, a machine is entirely object, and the subject has disappeared. At one point, Daston and Galison (2007) remark on the historical contingency of this division, "All history can do is to demonstrate the possibility of alternatives, thereby turning an apparent axiom – things that could never been otherwise than as we know them – into a matter for reasoned argument" (p. 376), yet, despite this admission of the constructedness of the human/machine binary (or the subject/object distinction in their case), they persist in believing in the accuracy of its split between two real types of phenomena. As Latour (2005) points out though, "There exists no relation whatsoever between 'the material' and 'the social world,'" or in this case the human and the machine, "because it is this very division which is a complete artifact" (p. 76). Thus, the error does not reside in the belief that, as Daston and Galison (2007) argue, the subject/object division is an arbitrary split, but rather in the belief that the division cleanly splits *anything at all*.

Imagining Machines

Besides being merely inattentive to the intricacies of these phenomena, the danger of this stubborn attachment to the human/machine binary from the beginning of any socio-historical analysis is that it obfuscates the hybridity and heterogeneity of observed phenomena and "prevents the understanding of a collective" (Latour 1999:180). The imposition of the human/machine binary is tantamount to blackboxing, a concept which, for Latour (1999), renders "the joint production of actors and artifacts entirely opaque" (p. 183) by reducing the complexity of the phenomenon at hand. Instead of beginning with any a priori ideas about who the various actors involved in any exchange are, Latour (2005) urges for a close tracing of the movements made. Using this Latourian analysis on machine learning in diagnostics, one concludes something very different from the subject/object, or human/machine, approach.

Narratives dependent on this absolute divide between human/machine will inevitably conclude that the development of machine learning techniques for diagnostics leads to the complete elimination of the subjective. Because the tasks that are typically performed by the subject are delegated to the automated machine, it is reasoned that almost complete (digital/mechanical) objectivity has been obtained. However, this account purposefully ignores this act of delegation, this translation of human tasks into machine tasks.

In machine learning techniques, a diagnostician first divides a given set of pre-labeled, "training" data into a set number of discrete groups. These data are then fed into the computer. The machine then analyzes and sorts them based on their likelihood of belonging in each of these discrete group, according to any number of parameters that are used to distinguish them (i.e., their features). The entire process generates an algorithm that can be later used for the separation of unlabeled data. So, for example, if an algorithm is "trained" to learn the difference between the MRI data from two individuals, one of which is already known to be depressed and the other healthy, then it can be applied to new data without pre-existing labels to sort them.

Without even getting into the computer design itself, it is apparent that there are numerous stages in which there is a trace of the human. Humans serve as gate-keepers for which data enters into the algorithm, the data are pre-labeled by an original diagnostician, and the number of groups is also divided by a human. Furthermore, often times the features selection is done by a human; even when it is not, in the case of unsupervised, data-driven methods, the overall algorithm is still a human construction. Other times, "simulated" or "virtual" data (in which a given data set is doubled and then used to develop the algorithm) is used as "training data," another decision influenced by humans. It is obvious that, by tracing the human and the non-human in the development of these techniques, there is not a simple division between human and non-human, but rather a complex fusion. Yet an account that begins with subject/object will only reveal the subjective human and the objective, nonhuman machine.

In the machine learning studies, an arbitrary division is made between the human diagnostician and the machine algorithm in order to authorize the knowledge produced by the machine. By conferring agency onto the machine (Latour 1988), the machine is allowed to "speak for itself"; in this manner, human acts of knowledge production can be delegated to and translated through the objective machine. Moreover, by then obfuscating the fusion of human elements within it, the machine can appear to produce universal, objective truths. This approach has been critiqued by a number of feminist scholars (Haraway 1988; Smith

1990). Haraway (1988), who argues that traditional, masculine accounts of objectivity, such as those espoused in philosophical epistemology of science (particularly those from the logical positivist school), cast this form of knowledge as bodiless: Existing entirely outside and independently of the human subject, objective knowledge does not reflect a finite and particular point of view. Rather, objective knowledge is believed to be universally truth. It does not have a history of being existent, only a history of being known. Once knowledge becomes labeled objective, it is elevated to a privileged status, which is irrefutable and uncontestable.

This research shows that such a formulation of knowledge is problematic. If the instantiations of objective knowledge cannot be challenged, they become a tool for subjugating those subjects with bodies, for those who necessarily speak from a finite, local, and particular point of view. By elevating the position of the machine to one of superiority, to one of pure numbers, the knowledge produced from the machine is made not only disembodied but, importantly, also more desirable than other knowledge: The machine is faster, impervious to error, and perceptible beyond the limitations of the human. Constructing this narrative of advantage and progress defines, "by contrast, an archaic and stable past. Furthermore, the word [progress] is always being thrown in the middle of a fight, in a quarrel where there are winners and losers" (Latour 1993:11).

Adopting Haraway's notion of "situated knowledge," I suggest we localize all knowledge within a body, to ground the disembodied specter of objective knowledge in a particular stance or viewpoint. If successful, this move would eliminate the possibility of the Archimedean point of view, the "conquering gaze from nowhere" (Haraway 1988:176), that which "mythically inscribes all the marked bodies, that makes the unmarked category claim the power to see and not be seen, to represent while escaping representation" (Haraway 1988:176). Situated knowledge possesses a necessary ethic of accountability because it identifies a view that can be located, questioned, and critiqued:

> We need to learn in our bodies, endowed with primate color and stereoscopic vision, how to attach the objective to our theoretical and political canners in order to name where we are and are not, in dimensions of mental and physical space we hardly know how to name. So, not so perversely, objectivity turns out to be about particular and specific embodiment, and definitely not about the false vision promising transcendence of all limits and responsibility. The moral is simple: only partial perspective promises objective vision. This is an objective vision that initiates, rather than closes off, the problem of responsibility for the generativity of all visual practices. Partial perspective can be held accountable for both its promising and its destructive monsters (Haraway 1988: 177).

Situated knowledge is always open to reexamination and reimagination. Machine learning discourse must be localized within the specific practices of human scientists to prevent the possibility of a corpus of knowledge about the brain-self devoid of participation from the people being diagnosed.

Conclusion

My inquiry into machine learning indicates that the machine itself is not pure object but an intricately-networked pastiche of human diagnostician and nonhuman algorithm, the complexity of which gets condensed by researchers through a series of rhetorical practices that erases any traces of the human and glorifies the abilities of the objective machine. Locating these traces of the human and destabilizing the machine as pure object, as my research has done, are necessary first steps to subjecting these discourses of power to participatory critique.

And in machine-based diagnosis, there also exists a cooperative hybrid of man and machine, in what I have proposed to call *cyborg diagnostics*. The machine neither replaces the human nor gets forgotten but is instead combined with the individual judgment of the clinician to reach the truth of a disease state. In both cases, the human and machine are tightly interwoven, neither deserving of more attention than the other. Previous STS work has extensively researched the practices of scientists. In addition, I propose to focus now on the *practices of machines* – to follow their movements and outline their developments in a narrative separate from that of unmediated technological progress.

As machine learning is being applied to other neuroimaging techniques, an expansion of the present to study to include other imaging modalities might be fruitful. In addition, expanding the analysis to add in non-human participants or even non-disease neuroimaging analyses (e.g., brain state prediction or brain mapping) would also prove relevant. If the technique continues to grow, research could focus on popular media outlets; more importantly, a clinical and/or research ethnography could be conducted to investigate how the scientists predictions and conclusions are implemented into practice and to begin to construct a tale of the lives of machines living among us.

Today, we are facing an increasing degree of automation in the scientific field. The proliferation of machines is part of a project to create an objective body of knowledge vis-à-vis the automaticity of scientific procedures and measurements. I urge future research to develop a theory of objectivity that takes the hybridity of the machine as a starting point. Only then can a rigorous and complete history of the present be undertaken.

References

Anderson, A., Dinov, I.D., Sherin, J.E., Quintana, J., Yuille, A.L., and M.S. Cohen. 2010. “Classification of spatially unaligned fMRI scans.” *Neuroimage* 49: 2509-2519.

Anderson, A., Labus, J.S., Vianna, E.P., Mayer, E.A., and M.S. Cohen. 2010. “Common component classification: What can we learn from machine learning?” *Neuroimage* [epub ahead of print].

Ardekani, B.A., Tabesh, A., Sevy, S., Robinson, D.G., Bilder, R.M., and P.R. Szeszko. 2011. “Diffusion Tensor Imaging Reliably Differentiates Patients with Schizophrenia from Healthy Volunteers.” *Human Brain Mapping* 32: 1-9.

Ashburner, J. and S. Kloppel. 2010. “Multivariate models of inter-subject anatomical variability.” *Neuroimage* [epub ahead of print].

Avants, B.B., Cook, P.A., Ungar, L., Gee, J.C., and M. Grossman. 2010. “Dementia induces correlated reductions in white matter integrity and cortical thickness: A multivariate neuroimaging study with sparse canonical correlation analysis.” *Neuroimage* 50: 1004-1016.

Beaulieu, Anne. 2001. “Voxels in the Brain: Neuroscience, Informatics and Changing Notions of Objectivity.” *Social Studies of Science*. 31(5): 635-80.

Blanchet, L., Krooshof, P.W.T., Postma, G.J., Idema, A.I., Goraj, B., Heerschap, A., and Buydens, L.M.C. “Discrimination between Metastasis and Glioblastoma Multiforme Based on Morphometric Analysis of MR Images.” *American Journal of Neuroradiology* 32: 67 – 73.

Brickman, A.M., Habeck, C., Ramos, M.A., Scarmeas, N., and Y. Stern. 2008. “A Forward Application of Age Associated Gray and White Matter Networks.” *Human Brain*

Mapping 29: 1139-1146.

Burge, J., Lane, T., Link, H., Qiu, S., and V.P. Clark. 2009. "Discrete Dynamic Bayesian Network Analysis of fMRI Data." *Human Brain Mapping* 30: 122-137.

Chen, R. and E.H. Herskovits. 2010. "Machine-learning techniques for building a diagnostic model for very mild dementia." *Neuroimage* 52: 234-244.

Chou, Y-Y., Lepore, N., Avedissian, C., Madsen, S.K., Parikshak, N., Hua, X., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W., Toga, A.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Mapping correlations between ventricular expansion and CSF amyloid and tau biomarkers in 240 subjects with Alzheimer's disease, mild cognitive impairment and elderly controls." *Neuroimage* 46: 394-410.

Costafreda, S.G., Dinov, I.D., Tu, Z., Shi, Y., Liu, C-Y., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Wahlund, L-O., Spenger, C., Toga, A.W., Lovestone, S., and A. Simmons. 2011. "Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment." *Neuroimage* [epub ahead of print].

Craddock, Cameron, Paul Holtzheimer, Xiaoping Hu, and Helen Mayberg. 2009. "Disease State Prediction from Resting State Functional Connectivity." *Magnetic Resonance in Medicine* 62: 1619-28.

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M-O., Chupin, M., Benali, H., Colliot, O., and the Alzheimer's Disease Neuroimaging Initiative. 2010. "Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database." *Neuroimage* [epub ahead of print].

Daston, Lorraine and Peter Galison. 2007. *Objectivity*. New York: Zone Books.

Davatzikos, C., Resnick, S.M., Wu, X., Parmpi, P., and C.M. Clark. 2008. "Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI." *Neuroimage* 41: 1220-1227.

De Boer, R., Schaap, M., van der Lijn, F., Vrooman, H.A., de Groot, M., van der Lugt, A., Ikram, M.A., Vernooij, M.W., Breteler, M.M.B., and W.J. Niessen. 2011. "Statistical analysis of minimum cost path based structural brain connectivity." *Neuroimage* 55: 557-565.

Duchesne, S., Bernasconi, N., Bernasconi, A., and D.L. Collins. 2006. "MR-based neurological disease classification methodology: Application to lateralization of seizure focus in temporal lobe epilepsy." *Neuroimage* 29: 557-566.

Dumit, Joseph. 2004. *Picturing Personhood: Brain Scans and Biomedical Identity*. Princeton, NJ: Princeton University Press.

Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Mirando, J., Marquand, A., Daly, E.M., Brammer, M.J., Murphy, C., Murphy, D.G., and the MRC AIMS Consortium. 2010. "Investigating the predictive value of whole-brain structural MR scans in autism: A pattern classification approach." *Neuroimage* 49:44-56.

Emblem, K.E., Zoellner, F.G., Tennoe, B., Nedregaard, B., Nome, T., Due-Tonnessen, P., Hald, J.K., Schele, D., and A. Bjornerud. 2008. "Predictive Modeling in Glioma Grading From MR Perfusion Images Using Support Vector Machines." *Magnetic Resonance in Medicine* 60: 945-952.

Emblem, K.E., Nedregaard, B., Hald, J.K., Nome, T., Due-Tonnessen, P., and A. Bjornerud. 2009. *Journal of Magnetic Resonance Imaging* 30: 1-10.

Fan, Y., Resnick, S.M., Wu, X., and C. Davatzikos. 2008. “Structural and functional biomarkers of prodromal Alzheimer’s disease: A high-dimensional pattern classification study.” *Neuroimage* 41: 277-285.

Ferranini, L., Palm, W.M., Olofsen, H., van der Landen, R., van Buchem, M.A., Reiber, J.H.C., and F. Admiraal-Behloul. 2008. “Ventricular Shape Biomarkers for Alzheimer’s Disease in Clinical MR Images.” *Magnetic Resonance in Medicine* 59: 260-267.

Filopoyv, R. Davatzikos, C., and the Alzheimer’s Disease Neuroimaging Initiative. 2011. “Semi-supervised pattern classification of medical images: Application to mild cognitive impairment (MCI).” *Neuroimage* [epub ahead of print].

Focke, N.K., Helms, G., Scheewe, S., Pantel, P.M., Bachmann, C.G., Dechent, P., Ebentheuer, J., Mohr, A., Paulus, W., and C. Trenkwalder. 2011. “Individual Voxel-Based Subtype Prediction Can Differentiate Progressive Supranuclear Palsy From Idiopathic Parkinson Syndrome and Healthy Controls.” *Human Brain Mapping* [epub ahead of print].

Franke. K., Ziegler, G., Kloppel, S., Gaser, C., and the Alzheimer’s Disease Neuroimaging Initiative. 2010. “Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters.” *Neuroimage* 50: 883-892.

Georgiadis, P., Cavouras, D., Kalatzis, I., Glotsos, D., Athanasiadis, E., Kostopoulos, S., Sifaki, K., Malamas, M., Nikiforidis, G., and E. Solomou. 2009. “Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textual features and ensemble pattern recognition methods." *Magnetic Resonance Imaging* 27: 120-130.

Georgiadis, P., Kostopoulos, S., Cavouras, D., Glotsos, D., Kalatzis, J., Sifaki, K., Malamas, M., Solomou, E., and G. Nikiforidis. 2011. "Quantitative combination of volumetric MR imaging and MR spectroscopy data for the discrimination of meningiomas from metastatic brain tumors by means of pattern recognition." *Magnetic Resonance Imaging* [epub ahead of print].

Gerardin, E., Chetelat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H-S., Neithammer, M., Dubois, B., Lehericy, S., Garnero, L., Eustache, F., Colliot, O., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging." *Neuroimage* 47: 1476-1486.

Gibson, E., Gao, F., Black, S.E., and N.J. Lobaugh. 2010. "Automatic Segmentation of White Matter Hyperintensities in the Elderly Using FLAIR Images at 3T." *Journal of Magnetic Resonance Imaging* 31: 1311-22.

Gong, Q., Wu, Q., Scarpazza, C., Lui, S., Jia, Z., Marquand, A., Huang, X., McGuire, P., and A. Mechelli. 2011. "Prognostic prediction of therapeutic response in depression using high-field MR imaging." *Neuroimage* [epub ahead of print].

Guo, Y., Bowman, F.D., and C. Kilts. 2008. "Predicting the Brain response to Treatment Using a Bayesian Hierarchical Model With Application to a Study of Schizophrenia." *Human Brain Mapping* 29: 1092-1109.

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." Pp. 172-188 in *The Science Studies Reader*, edited by Mario Biagioli. New York: Routledge.

Hinrichs, C., Singh, V., Mukherjee, L., Xu G., Chung, M.K., Johnson, S.C., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset." *Neuroimage* 48: 138-149.

Hinrichs, C., Singh, V., Xu, G., Johnson, S.C., and the Alzheimer's Disease Neuroimaging Initiative. 2011. "Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population." *Neuroimage* 55: 574-589.

Hu, X., Wong, K.K., Young, G.S., Guo, L., and S.T. Wong. 2011. "Support Vector Machine Multiparametric MRI Identification of Pseudoprogression From Tumor Recurrence in Patients with Resected Glioblastoma." *Journal of Magnetic Resonance Imaging* 33: 296-305.

Hua, X., Lee, S., Yanovsky, I., Leow, A.D., Chou, Y-Y., Ho, A.J., Gutman, B., Toga, A.W., Jack Jr., C.R., Bernstein, M.A., Reiman, E.M., Harvey, D.J., Kornak, J., Schuff, N., Alexander, G.E., Weiner, MW., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: An ADNI study of 515 subjects." *Neuroimage* 48: 668-681.

Hua, X., Lee, S., Hibar, D.P., Yanovsky, I., Leow, A.D., Toga, A.W., Jack Jr., C.R., Bernstein, M.A., Reiman, E.M., Harvey, D.J., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative. 2010. "Mapping Alzheimer's disease progression in 1309 MRI scans: Power estimates for different inter-scan intervals." *Neuroimage* 51: 63-75.

Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z., and E.H. Herskovits. 2010. "Predictive models of autism spectrum disorder based on brain regional cortical thickness." *Neuroimage* 50: 589-599.

Joyce, Kelly. 2008. *Magnetic Appeal: MRI and the Myth of Transparency*. Ithaca, NY: Cornell University Press.

Juntu, J., Sijbers, J., De Backer, S., Rajan, J., and Dirk Van Dyck. 2010. "Machine Learning Study of Several Classifiers Trained With Texture Analysis Features to Differentiate Benign from Malignant Soft-Tissue Tumors in T1-MRI Images." *Journal of Magnetic Resonance Imaging* 31: 680-689.

Kassner, A. and R.E. Thornhill. 2010. "Texture Analysis: A Review of Neurologic MR Imaging Applications." *American Journal of Neuroradiology* 31:809-816.

Kasparek, T., Thomaz, C.E., Sato, J.R., Schwarz, D., Janousova, E., Marecek, R., Rikryl, R., Vanicek, J., Fujita, A., Ceskova, E. 2011. "Maximum uncertainty linear discrimination analysis of first-episode schizophrenia subjects." *Psychiatry Research: Neuroimaging* [epub ahead of print].

Kodipaka, S., Vemuri, B.C., Rangarajan, A., Leonard, C.M., Schmallfus, I., and S. Eisenschenk. 2007. "Kernel Fisher discriminant for shape-based classification." *Medical Image Analysis* 11: 79-90.

Koutsouleris, Nikolaos, Eva M. Meisenzahl, Christos Zavatzikos, Ronald Bottlender, Thomas Frodl, Johanna Scheuerecker, Gisela Schmitt, Thomas Zetzsche, Petra Decker, Maximilian Resier, Hans-Jurgen Moller, and Christian Gaser. 2009. "Use of Neuroanatomical Pattern Classification to Identify Subjects in At-Risk Mental States of Psychosis and Predict Disease Transition." *Archives of General Psychiatry* 66(7): 700-712.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.

------. 1988. *The Pasteurization of France*. Cambridge, MA: Harvard University Press.

------. 1993. *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.

------. 1999. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

------. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York: Oxford University Press.

Lawyer, G., Nesvag, R., Varnas, K., Frigessi, A., and I. Agartz. 2008. "Investigating possible subtypes of schizophrenia patients and controls based on brain cortical thickness." *Psychiatry Research: Neuroimaging* 164: 254-264.

Lemm, S., Blankertz, B., Dickhaus, T., and K-R. Muller. 2011. "Introduction to machine learning for brain imaging." *Neuroimage* [epub ahead of print].

Leung, K.K., Barnes, J., Ridgway, G.R., Bartlett, J.W., Clarkson, M.J., Macdonald, K., Schuff, N., Fox, N.C., Ourselin, S., and the Alzheimer's Disease Neuroimaging Initiative. 2010. "Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease." *Neuroimage* 51: 1345-1359.

Li, S., Shi, F., Pu, F., Li, X., Jiang, T., Xie, S., and Wang, Y. 2007. "Hippocampal Shape

Analysis of Alzheimer Disease Based on Machine Learning Methods." *American Journal of Neuroradiology* 28: 1339-45.

Lin, F., Yu, C., Jiang, T., Li, K., Zhu, C., Zhu, W., Qin, W., Duan, Y., Xuan, Y., Sun, H., and P. Chan. 2006. "Discriminative analysis of relapsing neuromyelitis optica and relapsing-remitting multiple sclerosis based on two-dimensional histogram from diffusion tensor imaging." *Neuroimage* 31: 543-549.

Magnin, B., Mesrob, L., Kinkingnehun, S., Pelegrini-Issac, M., Colliot, O., Sarazin, M., Dubois, B., Lehericy, S., and Benali, H. 2009. "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI." *Neuroradiology* 51: 73-83.

Misra, C., Fan, Y., and C. Davatzikos. 2009. "Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI." *Neuroimage* 44: 1415-1422.

Morey, R.A., Petty, C.M., Xu, Y., Hayes, J.P., Wagner II, H.R., Lewis, D.V., LaBar, K.S., Styner, M., and G. McCarthy. 2009. "A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes." *Neuroimage* 45: 855-866.

Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Avedissian, C., Madsen, S.K., Parikshak, N., Hua, X., Toga, A.W., Jack Jr., G.R., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative. 2008a. "Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls." *Neuroimage* 43: 59-68.

Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Avedissian, C., Madsen, S.K., Parikshak, N., Toga, A.W., Jack Jr., C.R., Schuff, N., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls." *Neuroimage* 45: S3-S15.

Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Avedissian, C., Madsen, S.K., Parikshak, N., Hua, X., Toga, A.W., Jack Jr., C.R., Schuff, N., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Automated 3D Mapping of Hippocampal Atrophy and Its Clinical Correlates in 400 Subjects with Alzheimer's Disease, Mild Cognitive Impairment, and Elderly Controls." *Human Brain Mapping* 30: 2766-2788.

Mortamet, B., Bernstein, M.A., Jack, C.R., Gunter, J.L., Ward, C., Britson, P.J., Meuli, R., Thiran, J.-P., Kruegner, G., and the Alzheimer's Disease Neuroimaging Initiative. 2009. "Automatic Quality Assessment in Structural Brain Magnetic Resonance Imaging." *Magnetic Resonance in Medicine* 62: 365-372.

Nouretdinov, I., Costafreda, S.G., Gammerman, A., Chervonenkis, A., Vovk, V., Vapnik, V., and C.H.Y. Fu. 2010. "Machine learning classification with confidence: Application of transducive conformal predictors to MRI-based diagnostic and prognostic markers in depression." *Neuroimage* [epub ahead of print].

Pereira, F., Mitchell, T., and M. Botvinick. 2009. "Machine learning classifiers and fMRI: A tutorial overview." *Neuroimage* 45: S199-S209.

Perlbarg, V., Puybasset, L., Tollard, E., Lehericy, S., Benali, H., and D. Galanaud. 2009.

"Relation Between Brain Lesion Location and Clinical Outcome in Patients with Severe Traumatic Brain Injury: A Diffusion Tensor Imaging Study Using Voxel-Based Approaches." *Human Brain Mapping* 30: 3924-3933.

Phillips, C.L., Bruno, M-A., Maquet, P., Boly, M., Noirhomme, Q., Schnakers, C., Vanhaudenhuyse, A., Bonjean, M., Hustinx, R., Moonen, G., Luxen, A., and S. Laureys. 2010. "'Relevance vector machine' consciousness classifier applied to cerebral metabolism of vegetative and locked-in patients." *Neuroimage* [epub ahead of print].

Plant, C., Teipel, S.J., Oswald, A., Bohm, C., Meindl, T., Mourao-Miranda, J., Bokde, A.W., Hampel, H., and M. Ewers. 2010. "Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease." *Neuroimage* 50: 162-174.

Prasad, Amit. 2006. "Making Images/Making Bodies: Visibilizing and Disciplining through Magnetic Resonance Imaging (MRI)." *Science, Technology & Human Values* 30: 291-316.

Qing, L., Haiteng, J., Haiyan, L., Gang, L., Gaojun, T., and Y. Zhijian. 2010. "Depression Severity Evaluation for Female Patients Based on a Functional MRI Model." *Journal of Magnetic Resonance Imaging* 31: 1067-1074.

Radulescu, A.R., and L.R. Mujica-Parodi. 2009. "A principal component network analysis of prefrontal-limbic functional magnetic resonance imaging time series in schizophrenia patients and healthy controls." *Psychiatry Research: Neuroimaging* 174: 184-194.

Raj, A., Mueller, S.G., Young, K., Laxer, K.D., and M. Weiner. 2010. "Network-level analysis of cortical thickness of the epileptic brain." *Neuroimage* 52: 1302-1313.

Rizk-Jackson, A., Stoffers, D., Sheldon, S., Kuperman, J., Dale, A., Goldstein, J., Corey-

Bloom, J., Poldrack, R.A., and A.R. Aron. 2010. "Evaluating imaging biomarkers for neurodegeneration in pre-symptomatic Huntington's disease using machine learning techniques." *Neuroimage* [epub ahead of print].

Sato, J.R., Martin, M.d.G.M., Fujita, A., Mourao-Miranda, J., Brammer, M.J., and E. Amaro Jr. 2009. "An fMRI Normative Database for Connectivity Networks Using One-Class Support Vector Machines." *Human Brain Mapping* 30: 1068-1076.

Shen, H., Wang, L., Liu, Y., and D. Hu. 2010. "Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI." *Neuroimage* 49: 3110-3121.

Sikka, K., Sinha, N., Singh, P.K., and A.K. Mishra. 2009. "A fully automated algorithm under modified FCM framework for improved brain MR image segmentation." *Magnetic Resonance Imaging* 27: 994-1004.

Smith, Dorothy E. 1990. *The Conceptual Practices of Power: A Feminist Sociology of Knowledge*. Boston, MA: Northeastern University Press.

Stonington, C.M., Chu, C., Kloppel, S., Jack Jr., C.R., Ashburner, J., Frackowiak, R.S.J., and the Alzheimer's Disease Neuroimaging Initiative. 2010. "Predicting clinical scores from magnetic resonance scans in Alzheimer's disease." *Neuroimage* 51: 1405-1413.

Toews, M., Wells III, W., Collins, D.L., and T. Arbel. 2010. "Feature-based morphometry: Discovering group-related anatomical patterns." *Neuroimage* 49: 2318-2327.

Vemuri, P., Gunter, J.L., Senjem, M.L., Whitwell, J.L., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., and C.R. Jack Jr. 2008. "Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies." *Neuroimage* 39: 1186-1197.

Vemuri, P., Simon, G., Kantarci, K., Whitwell, J.L., Senjem, M.L., Przybelski, S.A., Gunter, J.L., Josephs, K.A., Knopman, D.S., Boeve, B.F., Ferman, T.J., Dickson, D.W., Pariski, J.E., Petersen, R.C., and C.R. Jack Jr. 2010. "Antemortem differential diagnosis of dementia pathology using structural MRI: Differential-STAND." *Neuroimage* 55: 522-531.

Wee, C-Y., Yap, P-T., Li, W., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., and D. Shen. 2011. "Enriched white matter connectivity networks for accurate identification of MCI patients." *Neuroimage* 54: 1812-1822.

Wen, W., Sachdev, P.S., Li, J.J., Chen, X., and K.J. Anstey. 2009. "White Matter Hyperintensities in the Forties: Their Prevalence and Topography in an Epidemiological Sample Aged 44-48." *Human Brain Mapping* 30:1155-1167.

Westman, E., Simmons, A., Zhang, Y., Muehlboeck, J-S., Tunnard, C., Liu, Y., Collins, L., Evans, A., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., Lovestone, S., Spenger, C., Wahlund, L-O., and the AddNeuroMed Consortium. 2011. "Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls." *Neuroimage* 54: 1178-1187.

Wilson, S.M., Ogar, J.M., Laluz, V., Growdon, M., Jang, J., Glenn, S., Miller, B.L., Weiner, M.W., and M.L. Gorno-Tempini. 2009. "Automated MRI-based classification of primary progressive aphasia variants." *Neuroimage* 47: 1558-1567.

Wu, M., Rosano, C., Butters, M., Whyte, E., Nable, M., Crooks, R., Meltzer, C.C., Reynolds III, C.F., and H.J. Aizenstein. 2006. "A fully automated method for quantifying and localizing white matter hyperintensities on MR images." *Psychiatry Research: Neuroimaging* 148: 133-142.

Yoon, U., Lee, J-M., Im, K., Shin, Y-W., Cho, B.H., Kim, I.Y., Kwon, J.S., and S.I. Kim. 2007. "Pattern classification using principal components of cortical thickness and its discriminative pattern in schizophrenia." *Neuroimage* 34: 1405-1415.

Young, K., Du, A.-T., Kramer, J., Rosen, H., Miller, B., Weiner, M., and N. Schuff. 2009. "Patterns of Structural Complexity in Alzheimer's Disease and Frontotemporal Dementia." *Human Brain Mapping* 30:1667-1677.

Zacharaki, E.I., Wang, S., Chawla, S., Yoo, D.S., Wolf, R., Melhelm, E.R., and C. Davatzikos. 2009. "Classification of Brain Tumor Type and Grade Using MRI Texture and Shape in a Machine Learning Scheme." *Magnetic Resonance in Medicine* 62: 1609-1618.

Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., and the Alzheimer's Disease Neuroimaging Initiative. "Multimodal classification of Alzheimer's disease and mild cognitive impairment." *Neuroimage* [epub ahead of print].

Zollner, F.G., Emblem, K.E., and L.R. Schad. 2010. "Support Vector Machines in DSC-Based Glioma Imaging: Suggestions for Optimal Characterization." *Magnetic Resonance in Medicine* 64: 1230-1236.