2019

Professional Judgment in an Era of Artificial Intelligence and Machine Learning

Frank A. Pasquale
University of Maryland Francis King Carey School of Law, fpasquale@law.umaryland.edu

Professional Judgment in an Era of Artificial Intelligence and Machine Learning

Frank Pasquale

Though artificial intelligence (AI) in healthcare and education now accomplishes diverse tasks, there are two features that tend to unite the information processing behind efforts to substitute it for professionals in these fields: reductionism and functionalism. True believers in substitutive automation tend to model work in human services by reducing the professional role to a set of behaviors initiated by some stimulus, which are intended to accomplish some predetermined goal, or maximize some measure of well-being. However, true professional judgment hinges on a way of knowing the world that is at odds with the epistemology of substitutive automation. Instead of reductionism, an encompassing holism is a hallmark of professional practice—an ability to integrate facts and values, the demands of the particular case and prerogatives of society, and the delicate balance between mission and margin. Any presently plausible vision of substituting AI for education and health-care professionals would necessitate a corrosive reductionism. The only way these sectors can progress is to maintain, at their core, autonomous professionals capable of carefully intermediating between technology and the patients it would help treat, or the students it would help learn.

Neoliberal ideology shapes both selves and society. The self is modeled as a maximizing or at least satisficing individual, seeking various forms of capital, power, and pleasure. Dominant evaluative modes are quantitative, algorithmic, and instrumentalist, focused on financialized rubrics of productivity (Beer, 2016). Society is a competition, primarily organized by markets, as designed and redesigned by state actors and, increasingly, by firms like Amazon and eBay (for their third-party sellers), TaskRabbit (for labor), Uber (for rides), and Google (for advertising) (Van Loo 2016).

Professionals in two human services sectors—health care and education—have offered sustained and extensive (if often unsuccessful) resistance to this neoliberal ideology of substitutive automation. Each sector values certain practices, defining them as *constitutive of* the field, rather than as mere means to an end. Nonprofits have a powerful presence in each sector, balancing mission and margin. Good physical and mental health, and knowledge, are ends in themselves, not merely means to accomplish something else. Excellent medical and educational practice do not easily admit of quantitative measurement. Qualitative evaluation, and a humble willingness to recalibrate and risk-adjust quantitative data, is crucial.

These professional ideals are all too often (and ironically) discounted in professional schools, especially business and law schools, and among the managers who stand to profit by commoditizing service industries (Khurana 2010, 2016)). Too many thought leaders in each depict automation as a veritable force of nature, driven forward by unstoppable currents of economic change. They advance rhetorics of automation, AI, and big data to devalue certain forms of labor by characterizing them either as routinizable, or in need of rationalization via machine learning

(Greenfield 2017). Many law firm partners and "legal tech" consultants deride legal research and writing as a task that computers can automate, in part to justify lower wages for current associates (Pasquale and Cashwell 2015). Non-professional workers have been modeled as a fungible source of data to be replaced by machines and software once a critical mass of their daily tasks is computerizable. Now, according to neoliberal devotees of disruption theory, it is time to bring that logic to the service sector.

Neoliberal managerialists promote AI as hard-headed common sense: the obvious next step to improve quality and cut costs. However, their views rest on a contestable epistemology of automation, grounded in the susceptibility of future situations to be translated into identifiable factors that machines can recognize and optimally respond to. This epistemology persistently ignores or elides the meaning of human practices, while focusing on their results. It hypostatizes the "data-driven," while minimizing the all-too-human process of gathering, cleaning, and analyzing data. Each of these steps is simultaneously vital to the strong AI vision of automation and are occasions for the exercise of professional judgment if they are to be done at all well. Data gathering and evaluation are not some trivial routine to be quickly dispatched at the outset of automation. Rather, they are essential to the entire project of automation.

Advocates of disruptive AI tend to presume that better technology (such as internet of things devices, or ubiquitous sensor networks) will simply accumulate the data necessary for quantum leaps in AI capacity. However, data about either processes or outcomes is rarely a simple "given," automatically captured by sensors or even algorithmically "smart" cameras and microphones (Gitelman 2013; Kitchin 2104). Instead, it is contestable, as multiple data sources, interpretations, rankings, and ratings reveal in service sectors ranging from health to education.

A wide-ranging literature on algorithmic accountability, as well as an academic and corporate movement for fairness, accuracy, and transparency in machine learning (FATML), is now revealing the scope and depth of bias, inaccuracy, and opacity in data and algorithms commonly employed in AI. These are valuable and important contributions to academic and professional discourses. However, current enthusiasm for reforming and improving AI will not address the theoretical infirmities of a neoliberal managerialist project premised on replacing human services professionals with software and machines. Indeed, reformists may well share these methodological meta-biases.

Though AI systems are growing more complex (Domingos 2015), reductionism and functionalism are two fundamental features of the information processing behind efforts to substitute AI and robotics for professionals. Reductionism extends Taylorist "scientific management" from manufacturing into professional contexts. Functionalism conceives of each part of a social order conducing to the operation of a larger system. Each is a poor fit for human service sectors, given the inevitably political, values-based, and conflictual nature of key aspects of good practice within them.

Robotics and AI, including even advanced machine-learning systems, comprehend professions as jobs, jobs as tasks, and tasks as observation, information processing, and actuation. Though such strategies to divide labor are sensible in many industrial contexts, they ignore the irreducibly holistic assessments that are hallmarks of good judgment.

Functionalism is another hallmark of automative epistemology. As a way of apprehending patients or students, functionalism moels their malaise or ignorance as an impediment to the proper functioning of society. For functionalists, the mind and body can be treated as "black boxes"—there is no need to explain how a given pattern of exposure to lectures and reading resulted in, say, a certain pattern of responses to multiple choice questions, or a certain average wage after graduation. The critical issue was simply figuring out what was the optimal pattern of stimulus to guarantee the right results in the future, to ensure proper functioning of one part (the student) in the whole (the labor market).

Despite massive investments in it, substitutive automation has not had much success in health and education. Instead of reductionism, an encompassing holism is a hallmark of professional practice—an ability to integrate facts and values, the demands of the particular case and prerogatives of society, and the delicate balance between mission and margin. Functionalism fails on other grounds—it elides the inevitably political, contestable, and conflictual aspects of professional life. Functionalism is properly a theory of biological and ecological systems to the extent that all parts work together to maintain homeostasis in an organism, or a sustainable balance of predators and prey in a given environment. It is always an uneasy fit in human systems because there are plural human goals and values, and the satisfaction of some entails the frustration of others. A technocratic "solution" to the problem of education is imaginable if the school and university really is a mere handmaiden to the labor market, but once other aims of education (such as the civic, aesthetic, and cultural), there is no singular system for the educator to function within.

Functionalism and reductionism combine to promote manipulative and normatively impoverished ways of modeling human interactions. Interpretive social science offers a much richer way of discussing the ways in which education and medicine can go well or poorly. A Habermasian model of communicative (as opposed to strategic) action recognizes the importance of intersubjective understanding and agreement in the classroom and the clinic (as the legal doctrine of "informed consent" suggests). Neither quality education nor quality medical care are reducible to a series of predetermined (or even machine-learned) steps. Knowledge, skill, and ethics are inextricably intertwined (Pasquale 2015a).

This essay exposes the suspect philosophical foundations of leading efforts to automate the health and education sectors. Just as few contemporary neoliberal thinkers would directly acknowledge the neoliberal foundations of their perspectives, AI mavens who would automate the professions rarely describe themselves as reductionists and functionalists (Mirowski 2014). Nevertheless, their methods directly reflect those commitments (Head 2014). Naming them unlocks decades of critiques, which ought to be available to those now seeking to resist blunt, substitutive automation in their professions. I summarize those counter narratives as holistic judgment and conflict theory. Both will stand as cornerstones of a humane professionalism that takes the diversity of human aims and aspirations seriously.

.

Disruption as Reductionism

Few experts in AI predict the imminent replacement of professionals by computers. However, in the disruption theory so popular in both Silicon Valley and Wall Street, technology should largely replace, rather than help, existing workers—even in human services. The critical idea here is that software and robotics can do for the health and education sectors what it once did for manufacturing: drastically expand production and reduce employment while cutting costs.

Leading business theorists push for robotically standardized work as a key to future advances in productivity in the health and education sectors. As business schools have abandoned the idea of management itself as a profession (Khurana 2010, 2016), their thought leaders look to alternative conceptualizations of managers' roles. Some model the replacement of labor with machines as a straightforward task: record and simulate a worker's pattern of actions, and then develop algorithms for their mechanical replication. The decline of privacy and the rise of surveillance expands the scope of such Taylorist aspirations (Bogard 1996).

For example, Harvard Business School professor Clayton Christensen sees Lasik-surgery machines as a model for future medical innovation, building expertise into equipment rather than relying on professionals for it (Christensen et al. 2009: 323–34). From this perspective, most doctors, most of the time, are not exercising judgment, like an artist or designer; instead, they are simply trying to match a set of symptoms to an optimal treatment modality (22–24). For disrupters, that problem is ideally suited to standardized, “value-adding processes,” like assembly-line work. Tomes like *Disrupting Class* and *The Innovator’s Prescription* lay out a blueprint for revolutionizing education and health care, respectively (Christensen et al. 2009; Christensen et al. 2008).

For two decades, Christensen has advanced a sweeping account of “disruption” as an explanation of business history and as the key to its future. According to disruption theory, nimble competitors replace established firms by developing rival products for the bottom end of the market. Initially cheap and of poor quality, these rival products end up dominating markets. Christensen’s theory of disruptive innovation has electrified the management consultant class, and its influence extends far beyond business. Thought leaders now aim to disrupt government as well (Eggers et al. 2013). Christensen has told both hospital and university leaders to shake up their operations (Kleinke 2009; Eyring and Christensen 2011). His public statements suggest that implementing disruptive principles can improve virtually every facet of human existence. Why, he asks, buy a single painting for your apartment when digital gallerists can email your flat screen “a fresh piece of art” (Lambert 2014) every three weeks? Disruption has become a theory of everything, catapulting Christensen to guru status as scholar, consultant, and sage.

Nevertheless, serious academics question the validity and relevance of disruption theory. Historian Jill Lepore's devastating *New Yorker* profile portrayed Christensen as an academic lightweight who downplays evidence that large, stable companies can sustain their business models (Lepore 2014). Business researchers Andrew A. King and Baljir Baatartogtokh (2015) have strengthened Lepore's case (Goldstein 2015). As Lee Vinsel (2015) observes, they found "only 9 of 77 cases that Christensen used as examples of disruptive innovation *actually fit* the criteria of his own theory." Given these embarrassments, it may be time to consign "disruption" to the dustbin of stale management theory buzzwords. Yet it is difficult for mere academics to debunk theories of "disruptive innovation," because they are less attempts to describe the world than blueprints for remaking it.

For over a decade, business books have exhorted managers to be "supercrunchers"—numbers-obsessed quantifiers, quick to make important decisions as "data driven" as possible. There is an almost evangelical quality to this work, a passionate belief that older, intuition-driven decisions are a sinful relic of a fallen world. In *Machine Platform Crowd* (2017), MIT professors Andrew McAfee and Erik Brynjolfsson aim to formalize the successive canonizations of statistics, big data, AI, and machine learning into a consultant-friendly catechism of what smart business leaders should do today.

In earlier iterations of AI, researchers tried to reduce human expertise to a series of propositions, rules to be applied by an expert system. Although this approach can work well for very narrow applications, it is difficult to formalize human reactions and skills into a series of rules. Contemporary approaches to machine learning attempt to overcome that problem by rapidly iterating potential responses to problems and evaluating the success (or likelihood of success) of each.[1] With enough data and computing power, machine-learning experts can try multiple algorithms to optimize performance. McAfee and Brynjolfsson mention the difficult problem of managing the temperature of a server farm, and it is easy to see how a computer program could solve the problem second-by-second better than any human expert, because there are so many variables (airflow, temperature outside, computational intensity in various parts of the building, etc.) that need to be computed nearly instantaneously. Moreover, a cutting-edge system can experiment, shifting allocations of cooling effort among, say, fans, air conditioners, and other methods, or determining whether a relocation of computing activity (toward, say, colder walls in winter) might be more cost-effective than increasing airflow in areas prone to overheating.

Various machine-learning methods are now being developed by different schools of computer scientists. Basic pattern recognizers can map a classic response to a given situation. Evolutionary algorithms can spawn a large number of approaches to a problem, experiment with which works best, and deploy the best method in the future. Bayesian classifiers can weigh evidence about whether a given strategy is working or not, modeling causation along arcs connecting different nodes in a network. And some programs even compose approaches on the fly, coming up with the types of nonhuman intelligence that wowed informed commentators during the victory of AlphaGo, Google’s Go-playing AI program, against the reigning Go champion in 2016.

McAfee and Brynjolfsson contrast what they describe as machines’ implacable, objective data analysis with humans’ tendency to distraction and subjective judgments. But their case for machine learning is overstated—even self-contradictory. To suggest that software can be optimized to make better decisions than humans, they offer a series of examples to demonstrate weaknesses in human judgment. A sociology professor used a mathematical model to predict firms’ adherence to budget and timeliness of product delivery better than purchasing managers. A county's nonverbal IQ test included more minority children in a gifted program than a process centered around parent and teacher nominations. Law professors' simple, six-variable model predicted Supreme Court rulings for the 2002 term better than eighty-three prominent legal experts did. From examples like these, and a simple behavioral economics story about human susceptibility to instinctual rashness, McAfee and Brynjolfsson conclude, "The evidence is overwhelming that, whenever the option is available, relying on data and algorithms alone usually leads to better decisions and forecasts than relying on the judgment of even experienced and 'expert' humans."[2]

But where do the algorithms and data come from? Experienced and expert humans. As digital sociologist Karen Gregory has observed, big data is made of people (Gregory, 2014). In most work settings, persons develop the algorithms to parse data generated by persons. People are part of the "crowd" that McAfee and Brynjolfsson (following Clay Shirky's *Here Comes Everybody*) praise for supplying data and labor to so many machine-learning applications, ranging from spam detection to targeted ads. Sophisticated work in critical algorithm studies repeatedly emphasizes the intertwining of computational and human elements in decision-making.

Thought leaders like Christensen, McAfee, and Brynjolfsson are too prone to elevate the *computational* aspects of business practices and to devalue their precursors and context. Computational thinking "is the thought processes involved in formulating problems so their solutions can be represented as computational steps and algorithms" (Aho 2011). In a brief, oft-cited article, Jeannette Wing has characterized computational thinking as "using abstraction and decomposition when attacking a large complex task or designing a large complex system" (among many other things—her article offers a litany of the particular tasks that exemplify the "steps and algorithms" Aho describes) (Wing 2006). This decomposition of practice into smaller and smaller parts is also a hallmark of the division of labor in general and its Taylorist application to professions in more recent literature on automation.

There is little doubt that computational approaches can help inform both health professionals and teachers as they do their work. Data about drug-drug interactions can inform clinical decision support systems, for example—or information about particularly effective lessons may build up as schools engage in more data collection. Nevertheless, the computational approach cannot substitute for professional judgment in myriad situations, which themselves cannot be identified by an algorithm in advance. For example, while early courses for mental health professionals may describe a series of routinized steps to begin a psychotherapeutic encounter, mental health care is not like a chess game, where every conceivable series of moves and countermoves are ideally mapped out in advance by authoritative teachers, administrators, or managers, let alone computer programs or algorithms (Weizenbaum 1976).

Nor does the ostensibly more mysterious AI behind Google's recent victory with its AlphaGo program portend well here. Set aside, for a moment, the usual debates over the inexplicability of machine learning—usually marshaled by AI proponents to demonstrate that it is or will be as flexible and spontaneous as human minds and thus worthy of similar respect and consideration (Gunkel 2014). Algorithmic thinking is fundamentally an optimization problem—a way of achieving a given end state or condition (O'Neil 2016; Eubanks 2017; Pasquale 2015b). Is there general agreement on what the indicia of a "cure" are in a psychotherapeutic encounter? Or whether orthopedic patients should "learn to live with" a slight ache, try a drug, or perform physical therapy (Reid 2010)? Each of these conditions, and countless more, are complex problems informed by multiple factors, whose identification itself will often justifiably vary among providers, insurers, and health system administrators.

There has long been a movement in US health care for “evidence-based medicine,” which seeks to reduce variation in the treatment of certain conditions in order to promote optimal care (Wennberg et al. 2007). US doctors routinely engage in unnecessary procedures (Charlesworth et al. 2016), and there is some evidence that automated clinical decision support software and other interventions could help nudge them away from such decisions (Tilson et al. 2016). Nevertheless, just as guardrails on a highway do not (yet) justify driverless cars, so too is the ability of automated clinical decision support to identify and warn against potentially unnecessary care merely necessary, but not sufficient, to a project of widespread automation of clinical practice.

The fate of clinical practice guidelines (CPGs) should serve as a cautionary tale for medicine’s computationalist reductionists. In the 1960s, physician groups began developing guidelines for the treatment of common conditions, like arrhythmia or migraine headaches. Such CPGs soon multiplied, covering many more conditions. Varying CPGs developed among different sets of practitioners; for example, there are at least five CPGs for treatment of congestive heart failure. Litigants started to use CPGs as evidence in some malpractice cases—to prove, for example, that a doctor had failed to meet the standard of care.

Soon, even more CPGs developed. Doctors who usually worked for the malpractice defense bar began to develop CPGs that were capacious—that is, they characterized a very wide range of practice as acceptable in response to a given condition (Mehlman 2012). Those who usually worked for plaintiffs did the opposite and tried to establish stricter standards (Furrow 2011). Those patterns may seem like a mere distortion of the medical system by warped malpractice law. However, even if malpractice causes of action were abolished, good faith disagreements over the scope and applicability of the guidelines would remain. As Ludwig Wittgenstein observed, only the simplest rules can articulate within themselves all the circumstances of their application (Schauer 1990). The rise of shared decision-making guides for conditions ranging from cancer to back pain demonstrates the importance of patient input into the caregiving process. Even if physicians of the future can deploy a combination of learning health-care system data and precision medicine to develop optimal treatment plans according to accepted rubrics in many cases, there will always be preference-sensitive care scenarios where deliberation on the values of the patient, with the patient, is critical. Moreover, it is very difficult to identify such scenarios in advance, so extracting the physician from "routine" care encounters would sacrifice critical opportunities for identifying these more complex, and value-laden, episodes of diagnosis, treatment, and support.

Bad-faith distortions of CPGs are also quite easy to imagine. Pharmaceutical firms may develop CPGs elevating the importance of their own blockbuster drugs (or, more likely, fund patients' rights groups to do so) (Choudhry 2002). Insurance companies, wary of expensive interventions, may make "watchful waiting" a more salient option (Woolf 1998). A computationalist might hope that all these groups would come together with doctors, economists, and other experts to hammer out one great and comprehensive CPG for each condition, as the Internet Engineering Task Force (IETF) has come to consensus on solving so many problems in online connectivity (Froomkin 2003). The IETF, however, decided about problems for which we have well-known parameters for service quality. At present, we do not have any principled governance procedure for setting the metrics for any number of medical conditions—or for establishing a rule of recognition that would allow a robot to recognize when a condition entered the realm of conditions with "well-defined metrics of success," as opposed to far more difficult scenarios.

Consider, say, a Stage IV congestive heart failure patient with an aortic valve issue who can either choose open heart surgery with a 10 percent chance of living eleven more months or a likely death within three months. That is clearly a computationally intractable scenario—whatever is to be decided must be talked over by the patient, doctors, the patient's family, and perhaps others (such as counselors, ministers or rabbis, and an insurer). When does the decision become "obvious"? When it is a 5 percent chance of living nine months? A 2 percent chance of living five months? I would never choose the surgery at that point; however, other patients may have other priorities or values.

This example may seem cherry-picked or overstated. However, dilemmas are legion. Kenneth Kaitin, director of the Tufts Center for Drug Development, describes "a cancer drug that only extends life by three months but allows you to go home, versus one that extends it a year but you end up staying in the hospital for the year, versus one that delays the progression of the disease but your overall outcome is not any different" (Swetlitz 2017). "There's so much gray area," he rightly concludes, implicating exactly the kinds of conflicts of values and introspection that algorithms are ill suited to address.

There are also parallels to it at every level of severity of conditions. Those suffering from a sinus infection should weigh the relative balance of risk and benefits of taking an antibiotic. Medical practices of risk reduction also present recurrent need for *holistic* assessments of a patient's health, including hard judgments about what is relevant to the holistic assessment. Even a practice as innocuous as taking one 81 mg aspirin a day (to ward off heart attacks) has some potential side effects (such as stomach bleeding). A reductionist may solve the trade-off by assigning a cardiac risk score to patients and recommending the prophylactic aspirin ingestion only to those with a certain level of risk. However, should a doctor also consider the possibility of stroke prevention via aspirin in this scenario? How many other articles in the medical literature on daily aspirin ingestion should enter the doctor's consideration—or any overall risk score on aspirin use? If there are multiple risk scores, how does one choose the appropriate one to consult? Again, these are very difficult decisions that depend, in part, on one's degree of confidence in the validity of the studies—a question by no means straightforward to answer in our contemporary age of "bent science" and "bad pharma" (McGarity and Wagner 2008; Mirowski 2011.

In short, Peter Denning is correct to question the universality of computational thinking:

> Alan Perlis . . . claimed that everyone can benefit from learning computational thinking. Other luminaries have followed suit. However, this general claim has never been substantiated with empirical research.
>
> For example, it is reasonable to question whether computational thinking is of immediate use for professionals who do not design computations—for example, physicians, surgeons, psychologists, architects, artists, lawyers, ethicists, realtors, and more. Some of these professionals may become computational designers when they modify tools, for example by adding scripts to document searchers—but not everybody. It would be useful to see some studies of how essential computational thinking is in those professions. (Denning 2017)

Denning correctly calls for more empirical research—but even better would be an acknowledgment that medicine, law, education, and other fields are independent spheres of judgment whose experts would understandably seek support from AI applications but would not be replaced by them (Styhre 2013). As Will Davies has argued, "A profession that claimed jurisdiction over everything would no longer be a profession but a form of epistemological tyranny" (Davies 2017). When computer scientist Pedro Domingos openly quests for a "Master Algorithm" that "can derive all knowledge in the world—past, present, and future—from data," he is pursuing precisely that form of "epistemological tyranny" (Domingos 2015). Such an approach reduces individuals to "dividuals," mere data points to be processed by machines (Sadowski and Pasquale 2015).

The Power of Reductionist Metrics

Contrasted with algorithmic imperialism, a diversity of professions appears to be a far more democratic and distributed way of recognizing and cultivating expertise. However, the professional bargain—workers granted autonomy, in exchange for advanced education, a continued commitment to keep up with the cutting edge of their fields, and fiduciary duties to the clients they work for—has been under attack for decades. Both left- and right-wing critics have impugned the motives of licensing boards that have controlled access to certain types of jobs (Larson 1977; MacDonald 1995; Abbott 1988). For the Right, professional autonomy with respect to some aspects of one's work offends the market principle that the purchaser of the service has the authority to define it (Gellhorn 1976). Left-wing critics have suggested a potentially conspiratorial elite: "If the possessors of [specialized] knowledge can form themselves into a group, which can then begin to standardize and control the dissemination of the knowledge base and dominate the market in knowledge based services, they will then be in a position to enter into a regulatory bargain with the state. This will allow them to standardize and restrict access to their knowledge, to control their market and supervise the production of producers" (Larson 1977: 71).[3]

Skepticism of the health and education sectors also created an opening for alternative modes of organizing work—or assessing its value—to gain traction.

For computationalist proponents of automation, metrics are a polestar of accountability, appealing to diverse ideological formations. To conservatives, outstanding performance on a well-recognized metric offers a kind of currency to a professional to compensate for market imperfections. For example, while most doctors do not charge what individuals can bear for what they do but instead must negotiate for payments with third-party payers, they cannot capture the full market value of exceptional performance. Ranking as the best spine surgeon or endocrinologist not only confers prestige, it can also enable the highly ranked professional to attract a large clientele or to charge high fees to cash-only patients (Pasquale 2007b). Where such rich rewards are unavailable, payors (ranging from government to private insurers to patients themselves) may simply avoid low-ranked doctors. School districts have demoted, reduced the pay of, or even fired teachers who perform poorly according to an algorithmic scoring of their performance—or, more accurately, the performance of their students on certain standardized tests.

For some progressives, metric discipline of professionals appeals as a way of controlling the discretion of groups perceived to be powerful elites. It is easy to see why leftists might be suspicious of, say, gastroenterologists earning $500,000 a year and demand that they demonstrate objective measures of performance, beyond hours worked or cases seen. Metrics also animated a core part of the compensation of Accountable Care Organizations' (ACOs) under the Medicare Shared Savings Program in the Affordable Care Act (ACA). This complex regulatory mechanism set forth thirty-three quality measurements for ACOs; they could share in the cost savings they generated only if they maintained or improved their scores on these metrics (Pasquale 2012). The Medicaid Authorization and CHIP Reauthorization Act of 2015 (MACRA) set forth even more complex mechanisms for Advance Payment Models (APMs), including ACOs, and yet another program for physicians who have not joined APMs.

Neoliberal dominance of the Democratic Party has also tempted its education policy leaders to impose more metrics on teachers, despite recurrent critiques of their reliability (O'Neil 2016) and effect on minority communities (Glynn and Waldeck 2013). Schoolteachers may not seem like much of an elite, as they earn less than similarly educated persons. Nevertheless, in a country where only about 33 percent of adults have a college degree, their pay is often above the median income, making them a target for the discipline of metrics.

The most effective way to storm the citadel of professional power is to question the distinction between expert and amateur. Critics of professions tend to see licensing as a hopelessly crude and unfair way of dividing the labor force into those capable of, say, practicing medicine or teaching, and those not qualified to do so. Metrics derived from big data are one way of dissolving the binary into a spectrum, from the most to the best qualified for any position.

To underpin this spectrum of assessment, big data and other forms of popular assessment could empower consumers to make their own judgments as to price/quality trade-offs instead of relying on professional boards. Teacher and doctor ratings, for example, could displace extant licensing and certification in those professions (Pasquale 2010b). According to the usual economic logic, such tiered rating (rather than all-or-nothing licensure) of professionals would expand access, gifting the poor the chance to pay far less for, say, health or education, by freely choosing lower-rated, cheap workers (including complete novices) in both fields. If the cheapest provider is a robot, virtual charter school, or app, all the better, from this perspective.

It is, however, often easy for the cynical or hypercompetitive to game rating and ranking systems (Pasquale 2011). Firms with substantial marketing budgets can invest in search engine optimization, review manipulation, and “astroturfed” recommendations (i.e., fake grassroots) (Reagle 2014). Money invested in influencing an intermediary’s ranking may be more important than actual performance; that is one reason why so many for-profits spend more on marketing than on instruction (Collini 2013). Corporations or foundations may pay think tanks to smear competitors with biased and one-sided reports. Competition can also reshape the very values it advances, by encoding (and, all too often, distorting) them in crude metrics (Pasquale 2007a).

Of course, we all dream of someday receiving proper recognition for our work, and sometimes this dream takes the form of hoping or demanding a ranking of its value. Managers recognize that they can take advantage of this altogether justifiable longing for recognition and direct it to their own purposes. The commensurating power of numbers, sweeping aside contestable narratives, promises a simple rank ordering of merit, whether in schools, hospitals, or beyond. Measurements are not simply imposed from the top down (Espeland and Sauder 2016).[4] They also colonize our own understandings of merit—and at the limit can count on self-interestedly rational support from the 49.9 percent of persons who are going to end up “above-average” on any percentile metric.

Those who do well on metrics should be very careful about promoting such measures, because they so often distort the social practice that they ostensibly measure (Edwards and Roy 2017). Measured on their thirty-day mortality rate (i.e., the percentage of patients who are still living thirty days after an operation), surgeons may simply avoid taking care of the very sick (Knapton 2016).[5] Cheating scandals have rocked schools ranked and rated on test scores. Even when the schools play fair, they may drop physical education, art, music, and other classes in order to teach to tests dominated by quantitative and verbal measures. As sociologist Donald T. Campbell put it in the eponymous Campbell’s Law, “The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor" (1975: 35).

Of course, debunkers of metrics should practice their exposure of gaming with discretion and humility. As organizers of a conference on gamed metrics at the University of California, Davis asked, "Can we reliably draw a clear separation between gaming the metrics game and engaging in misconduct?" (Innovating Communication in Scholarship 2015). Even if such distinctions are hard to make, pursuing them is a worthy task wherever metrics have a serious impact on resource allocation. They help expose the manipulability of supposedly objective measures of success.

Perhaps to avoid such manipulation, or to hide it, some institutions are now contracting with firms that boast of proprietary, secret algorithms designed to rank and rate employees. Rated individuals have a more difficult time gaming a metric when they cannot fully understand how it works (Pasquale 2011). Such measures tend to alienate knowledge workers. For example, the Rutgers Graduate Faculty voted to condemn such algorithmic ranking and rating by a 114–2 vote, rejecting out of hand black box and proprietary metrics (Flaherty 2016; Rutgers Graduate School Faculty 2016). They also demanded their data profiles. This was an important action based on documented concerns about the negative effects of black box algorithms in other sectors. As the activist mathematician Cathy O'Neil (2016) shows in her book *Weapons of Math Destruction*, algorithmic assessments of quality have unfairly denied critical opportunities for employment, career advancement, health, credit, and education. They deserve far more scrutiny than is common presently (Muller, 2018).

There is also a fair amount of hypocrisy in the deployment of such measures. Managers now use algorithmic assessment tools to sort employees worldwide on criteria of cost-effectiveness but spare themselves the same invasive surveillance and ranking (Peck 2013; Ashbrook 2013). Professionals in health and education should demand that administrators impose upon themselves similar reportable metrics of productivity—and rank themselves just as pervasively as they rank subordinates.

Even flagship journals of the technocratic policy apparatus have acknowledged rankings' fallibility. For example, an article in *Health Affairs* recently demonstrated that hospital rankings vary wildly, based on metrics like risk adjustment (e.g., how much an at-risk patient population should excuse poor outcome measures from a doctor or hospital) (Austin et al. 2015). There are always new "risks" (and "benefits") being discovered as influencers of health outcomes. Consider the controversy over the "epidemiological paradox" of the higher-than-expected health status of Latinos in the southwestern United States: despite worse secondary indicators of health (like blood pressure or obesity rates), this population appeared to live longer than many other groups with better numerical indicators of health (Markides and Coreil 1986).[6] A simple application of that fact to metrics of hospital performance would require us to "risk adjust" for ethnicity—that is, to carefully avoid giving too much credit to hospitals with a highly Latino patient base because their results are being boosted by that demographic mix. However, there are many explanations for the Latino Paradox, each of which prevails to varying degrees among the demographic mix at any given hospital. Do risk adjusters dive in to that granular of an assessment of patient mix? When do they stop the chain of risk adjustment (say, boosting the score of a hospital that takes on more patients with high blood pressure than with other conditions) and adjustments to risk adjustments (knocking the score back down a bit when it turns out the hospital has a high proportion of Latino patients)?

These are deeply political and philosophical—not just technological—questions. The authors of the *Health Affairs* study mentioned above concluded that hospital rankings should be fine-tuned to be ever-better indicators of the true quality of services provided. But what if a bad ranking decreases a hospital's number of privately insured patients (the most lucrative payers), reducing its resources, which in turn reduces its ability to do better in future rankings? A musical chairs logic of elimination could make sense for consumer goods that are discretionary purchases. When, however, it leads to the closure or weakening of "failing" hospitals and schools concentrated in poorer areas (a designation that can easily become a self-fulfilling prophecy), those left behind often must struggle to reassemble routine care.

In mental health care, the question of "outcome measurement" is usually more fraught. Just as there is an ideological battle over the proper scope or intensity of use of "cognition-enhancing" drugs, there is inevitably social disagreement over which aspects of life are necessary or sufficient to happiness, or even the alleviation of misery (Pasquale 2010a; Crary 2013). Introspection is not susceptible to multiple choice or true/false questions, however eagerly many "wellness vendors" may ask insured individuals to rate their marriage, job, community, and more, on a scale of 1 to 10 (ranging from "not at all satisfied" to "very satisfied") (Hull and Pasquale 2017). The extent and nature of obligations to family and community are contestable, and even if some practices of the self might improve well-being, it is not at all clear that an antiseptic form, often little more than a computerized confessional, is the proper place to prompt them or measure their impact. As David Morgan has observed, "Nothing human can be understood in the abstract. We also have to interpret thoughts, emotions, and behaviours in a social context" (Morgan 2016).

Three Kinds of Functionalism

Fundamentally *definitional* issues lie at the heart of the movement to automate professions—for example, whether a series of online sessions, tests, and "badges" count as a college degree, or a series of texts can count as a psychotherapeutic encounter. Metrics-focused reductionism feeds into functionalist frames here. Can the well-badged student get a job? Is app therapy helping users get out of a funk and back to productive employment? These approaches assume that there is a clear and simple *function* for education and health interventions, and that once an AI replacement for a human instructor or health-care provider performs that function, the human is effectively replaceable.

The term *functionalism* has many valences that are instructive here. For example, much of AI itself rests on a functionalist theory of mind—an idea that, "in principle, a machine (say, one of Isaac Asimov's robots), a human being, a creature with a silicon chemistry, and a disembodied spirit could all work much the same way when described at the relevant level of abstraction, and that it is just wrong to think that the essence of our minds is our 'hardware'" (Putnam 1988: xii). Philosopher Jerry Fodor has flatly stated that "thinking is computation" (Fodor 1998: 9).[7] To the extent the former activity is exhausted by the latter, each advance in computing technology is a step toward the replacement of brain-based "wetware" with in silico processing capacities. On this view, the articulation of the functions of any social role should generate a finite list of tasks (and algorithms for choosing which tasks to perform, when, in response to a finite list of stimuli). Rather than viewing human embodiment as *constitutive* of the social role of doctor or teacher, the functionalist sees the human as just one of many possible entities to perform such tasks.

Functionalism has also inspired a great deal of social science research. In both anthropology and sociology, functionalists have modeled social relations via an "organic analogy," which "compared the different parts of a society to the organs of a living organism" (Porth et al.

n.d.).[8] While the function of the stomach may be to digest, and the foot to aid in locomotion, the function of the education system may be modeled as preparation of workers for jobs, and the health-care system as keeping them healthy enough to perform productive labor (whether paid or unpaid). Of course, given what we now know about the many roles of gut bacteria in varied areas of health, the organic analogy appears crude from the start. Still, it is not in principle impossible to imagine various technologies taking over some roles in the body (such as a robotic pancreas), and similarly, the functionalist vision enables a way of comprehending society that is ultimately flexible and procedural, assuming a radical flexibility in the nature of institutions so long as they fulfill their necessary functions.

Finally, in US administrative and constitutional law, functionalism is a legal theory recommending flexibility in our treatment of the expanding power of various agencies of government. The US Constitution assigns judicial functions to a judiciary, legislative functions to Congress, and executive functions to the president and the thousands of agencies and subsidiary entities he or she leads. As agencies interpret and apply vague statutes, they issue rules (a quasi-legislative function) and adjudicate cases (a quasi-judicial function). Formalists have battled these manifestations of agency authority for decades, complaining that they grant outsized authority to the executive branch. Functionalists respond that there is simply no way of governing a nation as populous and diverse as the United States without delegating substantial power to agencies. Article III federal courts have tended to agree, realizing they have nowhere near the personnel necessary to replace, say, the hundreds of administrative law judges appointed pursuant to Article I to handle disability claims (let alone the thousands of other contexts where federal law is clarified and applied).

These three forms of functionalism—in philosophy, social science, and law—may appear, at first glance, to have little in common. Charles Taylor has cautioned that high-level models of self or society—whether reductionist or holistic—do not necessarily dictate any particular ideological or political position (Taylor 1989). However, there are elective affinities between the denigration of embodied humanness, the expansion of executive power, and the commercial/political pressure to reduce the power of professionals. As David Golumbia has explained in his genealogy of philosophical functionalism, computationalism joins biological reductionism in "the political philosophy we might call objectivism—the belief in a quasi-platonic 'world out there' that transcends the human social world" (2009: 78). In most automative, substitutive visions of replacing human professionals with some combination of apps, software, and robots, there is a facile assumption that some authority could tote up the quantitative metrics of success in a field and decree when a machine had met or excelled human performance in these metrics. Quantitative technocracy in the executive branch—such as the "regleprudence" of the Office of Management and Budget—is a far better fit for such a vision than the agonistic haggling of rival interests in the legislative branch (Davidson and Leib 2015).

Political deliberation, rhetoric, and logrolling in Congress recall one recurrent "other" of functionalist social theory—conflict theory. From Marxists (who emphasize class conflict) to pluralists (focused on rival interest groups), conflict theorists question the plausibility of models of social order operating harmoniously, like organs in a body. They emphasize the rivalrous, zero-sum nature of allocation of resources within political and organizational settings. To be sure, this realistic, perhaps even naturalistic, view does not exhaust political theory—currents of more idealistic deliberativism and social democracy animate much vital thinking about democracy. However, conflict theory does give us a sense of the inevitably contested nature of topics ranging from quotidian quality metrics and to the highest aspirations of medical and educational institutions.

Nor are quantitative metrics the only rational way to resolve such conflicts. Organizations are inevitably political, and both rhetoric and persuasion justly play an important role in political life (Garsten 2006). It may be possible to lock students in a room with a sophisticated computer program each day and not let them out until they have mastered some set of interactions with the machine. Even if such a method achieved better test outcomes than traditional lectures and exercises, it should not be preferred to them. Rather than relying on force, or even its softer relations of incentives and intangible penalties, good teachers inspire, cajole, delight, and encourage (Bain 2004). In this practice, they model, for their charges, the kind of discussion, deliberation, and leadership that can make them effective citizens and workers.

The value of dialogue in psychotherapeutic encounters, as well as in educational settings, should also be evident. There is wisdom in investing in human-focused care and education teams, rather than dispatching these duties to smartphones and apps. Students and patients should resist pervasive appification because it is a short step to behaviorism. For example, it is easy to imagine a future health app synced with a Pavlok wristband to shock diabetics who fail to take insulin at a certain time, after a series of warnings. Similarly, education technology prompting students to respond in stereotyped ways that are mechanically recognizable as adequate, menaces opportunities for creativity. This method of forcing behavior modification reflects an outsourcing of agency to machines. However effective it might be, it represents a disturbing externalization of the will—an analogue, at the personal level, to the type of decision a polity might make to install an authoritarian leader (Elster 1979).

Routes to Autonomy in an Era of Data-Driven Performance Metrics

None of these points are meant to dismiss performance metrics or rankings altogether. Instead, policymakers should assure lasting and meaningful involvement of professionals in the processes meant authoritatively to judge their value. Co-governance of innovation, among extant professionals and technologists, is key.

Ideally, professionals play some role in helping to construct the ranking systems that are measuring their performance, with an eye to improving their work. In the mid-2000s, doctors in New York and Connecticut found that insurers were rating them based on obscure metrics of "cost-effectiveness" that often boiled down to how profitable their practice was for the insurer. Both sued, winning settlements that imposed a wide array of conditions on such rankings and ratings—including transparency as to their calculation and the ability to contest scores and data (Madison 2009). The suits rested on consumer protection rationales. Academics afflicted with similarly problematic rankings and ratings may want to study the ways doctors successfully contested such rankings.[9]

We should also promote a positive vision of professional autonomy based on narrative assessments. It is not enough for us to express dissatisfaction with the metricization of accomplishment. As citation counts proliferate, accumulating the ersatz currency of reputational quantifications threatens to overwhelm the real purpose of research—just as financialization has all too often undermined the productive functions of the economy. Traditional modes of assessment (including tenure letters and Festschrift tributes) are an alternative form of evaluation. Essays explaining the shape of one's career, and one's reasons for choosing a certain topic or method of research, are a type of self-evaluation that should become more popular among scholars at certain career milestones (like tenure, appointment to full professor or senior lecturer, and, say, every five or ten years thenceforward.) We need better, more narrative, midcareer assessments of the depth and breadth of scholarly contributions. Such qualitative modes of evaluation can be far richer than the quantification-driven metrics now ascendant in the academy.

In short, it "takes a theory to beat a theory," and an alternative method of explaining what we do and how we can do it in better or worse ways is necessary to displace the hegemony of rankings. Such narratives may impose their own disciplines and anxieties. At least they promise to relieve us from the fantasy that managers and bureaucrats can judge scholars, doctors, nurses, teachers, and workers of all kinds along preset, commensurating metrics.

There is another, parallel reason to insist on a subordination of automation and standardization in classrooms and hospitals to the supervision of local domain experts—those primarily trained in healthcare and education, and only secondarily, if at all, in coding and software development. The relative autonomy of doctors and teachers reflects a hard-earned trust and a responsibility for governance. To the extent that persons, rather than distant algorithmic assessors, are running and staffing institutions, there is always some residue of autonomy and power entrusted to them locally. The kind and degree of that autonomy will always be as much a political as an economic question.

Of course, this distribution of power, enabling certain workers to enjoy some level of autonomy and control over how they do their work, is always fragile and contestable. Sometimes professionals abuse the public trust, and their self-regulation needs to be monitored and regulated by the state. Nor should such respect and autonomy be reserved for professionals alone. Well-ordered societies may decide to slow down the pace of automation in almost any field, in order to assure more democratic governance of its implementation. The social theorist Hartmut Rosa has been dismissive of calls to "shape human affairs against the self-autonomizing forces of acceleration" because he says "it remains unclear who the political and institutional bearers of such a politics of deceleration could be, and also because of the unpredictable economic and social costs that such a forced resynchronization would engender" (Rosa 2013: 321). However, both professions and unions have served, and can continue to serve, as these "political and institutional bearers of such a politics of deceleration."[10] Contemporary social theory must engage with existing regulatory bureaucracy and civil society institutions, rather than rejecting them out of hand with vague concerns about "unpredictable economic and social costs."

Nor is all-pervasive quantification and metricization the ineluctable logic of economic progress. For true believers in metrics and standardization, problems with existing metrics are simply a prompt to improve metrics. However, no advocate for civically focused education should feel obliged to articulate quantifiable measures of schools' ability to preserve and maintain knowledge and culture, adapt to economic change, and prepare citizens for self-governance (Allen 2016). Moreover, even if schools were to take on that well-nigh Sisyphean task of quantification, they could not specify outcome measures or key performance indicators in detail without attending to community input and values. These values are bound to evolve over time, as well.

Any presently plausible vision of substituting AI for education and health-care professionals would necessitate a corrosive reductionism, premised on patients and students accepting services as "medical care" or "education" that are far inferior to what a skilled, reflective practitioner in either field could provide. The only way these sectors can progress is to maintain, at their core, a large (and likely growing) set of professionals capable of carefully intermediating between technology and the patients it would help treat, or the students it would help learn. The alternative is an epistemological tyranny all too likely to support a political one.

NOTES

1. As Erik Brynjolfsson puts it, "If you give them enough examples the machine-learning algorithms figure out the rules on their own. That's a real breakthrough. It overcomes what we call Polanyi's paradox. Michael Polanyi the Polymath and philosopher from the 1960s famously said 'We all know more than we can tell,' but with machine learning we don't have to be able to tell or explain what to do. We just have to show examples. That change is what's opening up so many new applications for machines and allowing it to do a whole set of things that previously only humans could do" (Brynjolfsson 2017). The suppressed question here, of course, is who decides which examples to show. *That* initial step is full of decisions saturated with politics and values, which are all too often laundered out of machine learning's immaculate predictions.

2. Ironically, the parade of examples the authors give for the superiority of "data-driven" decision-making are themselves no more than a *narrative* of computationalist supremacy. McAfee and Brynjolfsson give us no sense of the universe of studies available on the comparative advantage of computation over human decision-making, the applicability of these studies, or even whether their examples have been replicated or reanalyzed. Without grounding in such basic statistical concepts, their sweeping claims (one study on Supreme Court prediction is a clue to the future of the entire legal industry; one logistics model could eliminate vast swathes of human labor in that field) will ring hollow to anyone with even the slightest critical faculty.

3. For a response to Larson's approach, see Remus 2015.

4. Michael Sauder and Wendy Nelson Espeland (2009) discuss how administrators are tempted to "game the system" by focusing resources on certain metrics. They "define gaming as cynical efforts to manipulate the rankings data without addressing the underlying condition that is the

target of measurement. [For example,] some schools encourage underqualified applicants to apply to boost their selectivity statistics." New methodologies may also be a kind of preemptive gaming. For example, Paul Caron (2011) has found that "in every alternative ranking of law schools, the ranker's school ranks higher than it does under U.S. News [the dominant ranking method]."

5. A simple metric of comparison might be the "thirty-day mortality rate" (i.e., the number of patients who die within thirty days of surgery). But that is manipulable—a patient might be kept on mechanical ventilation postsurgery for thirty-one days to count as a survivor in the reporting period.

6. Works citing this paper put forward a surprising array of explanations for the paradox.

7. To be clear: Fodor's position is contested. I do not accept it as given but merely mention it here as a conceptualization of thought that is highly influential among devotees of strong AI and automation that substitutes machines for workers.

8. The organic analogy goes back at least to Plato, as Anthony Kronman (1993) explains in his discussion of the city and the soul in *The Lost Lawyer*.

9. Note also how lawyers failed, losing on First Amendment grounds, when they sued Avvo for rating them. Vrdolyak v. Avvo, Inc., 206 F. Supp. 3d 1384 (N.D. Ill. 2016).

10. For example, such institutions may resist calls to cut college to three or two years, or to continually channel demand for medical care into cheap apps, when it is clear that the longer time period or personal touch both serve the student's/patient's best interests and those of society as a whole. There is no necessary tension between producer and consumer interests here, particularly when both can agree to develop public financing mechanisms that provide universal access to high quality services (Vaheesan and Pasquale, 2018).

REFS

Abbott, Andrew. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.

Aho, Alfred V. 2011. "Computation and Computational Thinking." *Ubiquity* (January 2011). http://ubiquity.acm.org/article.cfm?id=1922682.

Allen, Danielle. 2016. *Education and Equality*. Chicago: University of Chicago Press.

Ashbrook, Tom. 2013. "The High-Tech Hiring Market of Today." *On Point*, December 3, 2013. http://onpoint.legacy.wbur.org/2013/12/03/high-tech-hiring-employers-big-data.

Austin, J. Matthew, Ashish K. Jha, Patrick S. Romano, Sara J. Singer, Timothy J. Vogus, Robert M. Wachter, and Peter J. Pronovost. 2015. "National Hospital Ratings Systems Share Few Common Scores and May Generate Confusion Instead of Clarity." *Health Affairs* 34, no. 3:423–30. doi: 10.1377/hlthaff.2014.0201.

Bain, Ken. 2004. *What the Best College Teachers Do*. Cambridge, MA: Harvard University Press.

Beer, David 2016. <u>Metric</u> <u>Power</u>. London: Palgrave MacMillan.

Bogard, William. 1996. *The Simulation of Surveillance: Hypercontrol in Telematic Societies*. New York: Cambridge University Press.

Brynjolffson, Erik. 2017. "How AI Is Already Changing Business (Interview)." *Harvard Business Review*, July 20, 2017. https://hbr.org/ideacast/2017/07/how-ai-is-already-changing-business.

Campbell, Donald T. 1975. "Assessing the Impact of Planned Social Change." In *Social Research and Public Policies*, edited by Gene M. Lyons, 3–45. Hanover, NH: University Press of New England.

Caron, Paul. 2011. "Size Matters: Thomas Cooley's 2011 Law School Rankings." *TaxProf Blog*, February 9, 2011. http://taxprof.typepad.com/taxprof_blog/2011/02/size-matters-.html.

Charlesworth, Christina J., T. H. Meath, A. L. Schwartz, and K. J. McConnell. 2016. "Comparison of Low-Value Care in Medicaid vs Commercially Insured Populations." *JAMA Internal Medicine* 176, no. 7: 998–1004.

Choudhry, Niteesh. 2002. "Relationships between Authors of Clinical Practice Guidelines and the Pharmaceutical Industry." *JAMA* 287, no. 5: 612–17.

Christensen, Clayton M., Jerome H. Grossman, and Jason Hwang. 2009. *The Innovator's Prescription: A Disruptive Solution for Health Care.* New York: McGraw-Hill.

Christensen, Clayton M., Michael B. Horn, Louis Caldera, and Louis Soares. 2011. Disrupting College: How Disruptive Innovation Can Deliver Quality and Affordability to Postsecondary Education. Washington, DC: Center for American Progress. https://www.americanprogress.org/wp-content/uploads/issues/2011/02/pdf/disrupting_college.pdf.

Christensen, Clayton M., Michael B. Horn, and Curtis W. Johnson. 2008. *Disrupting Class: How Disruptive Innovation Will Change the Way the World Learns*. New York: McGraw Hill.

Collini, Stefan. 2013. "Sold Out." *London Review of Books*, October 24, 2013. https://www.lrb.co.uk/v35/n20/stefan-collini/sold-out.

Crary, Jonathan. 2013. *24/7: Late Capitalism and the Ends of Sleep*. New York: Verso.

Davidson, Nestor M., and Ethan J. Leib. 2015. “Regleprudence.” *Georgetown Law Journal* 103: 259–316.

Davies, William. 2017. “Elite Power under Advanced Neoliberalism.” *Theory, Culture and Society* 34, no. 5–6: 227–50. DOI: 10.1177/0263276417715072.

Denning, Peter J. 2017. “Remaining Trouble Spots with Computational Thinking.” *Communications of the ACM* 60, no. 6: 33–39.

Edwards, Marc A., and Siddhartha Roy. 2017. “Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition.” *Environmental Engineering Science* 34, no. 1: 51–61.

Eggers, William D., Laura Baker, and Audrey Vaughn. 2013. *Public Sector, Disrupted: How Disruptive Innovation Can Help Government Achieve More for Less*. New York: Deloitte University Press.

Elster, Jon. 1979. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. New York: Cambridge University Press.

Espeland, Wendy Nelson, and Michael Sauder. 2016. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. New York: Russell Sage Foundation.

Eubanks, Virginia. 2017. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin’s Press.

Eyring, Henry J., and Clayton M. Christensen. 2011. *The Innovative University: Changing the DNA of Higher Education from the Inside Out*. San Francisco: Jossey-Bass.

Flaherty, Colleen. 2016. “Refusing to Be Measured.” *Inside Higher Ed*, May 11, 2016. https://www.insidehighered.com/news/2016/05/11/rutgers-graduate-school-faculty-takes-stand-against-academic-analytics.

Fodor, Jerry A. 1998. *Concepts: Where Cognitive Science Went Wrong*. New York: Oxford University Press.

Froomkin, A. Michael. 2003. "Habermas@Discourse.Net: Toward a Critical Theory of Cyberspace." *Harvard Law Review* 116, no. 3: 749–873.

Furrow, Barry. 2011. "Regulating Patient Safety: The Patient Protection and Affordable Care Act." *University of Pennsylvania Law Review* 159:1728–1779.

Garsten, Bryan. 2006. *Saving Persuasion: A Defense of Rhetoric and Judgment*. Cambridge, MA: Harvard University Press.

Gellhorn, Walter. 1976. "The Abuse of Occupational Licensing." *University of Chicago Law Review* 44, no. 1: 6–27.

Gitelman, Lisa, ed. 2013. *"Raw Data" Is an Oxymoron*. Cambridge, MA: MIT Press.

Glynn, Timothy P., and Sarah E. Waldeck. 2013. "Penalizing Diversity: How School Rankings Mislead the Market." *Journal of Law and Education* 42, no. 3: 417–500.

Goldstein, Evan. 2015. "The Undoing of Disruption." *Chronicle of Higher Education*, September 15, 2015. http://chronicle.com/article/The-Undoing-of-Disruption/233101.

Golumbia, David. 2009. *The Cultural Logic of Computation*. Cambridge, MA: Harvard University Press.

Greenfield, Adam. 2017. *Radical Technologies: The Design of Everyday Life*. New York: Verso.

Gunkel, David J. 2014. "A Vindication of the Rights of Machines." *Philosophy and Technology* 27, no. 1: 113–32.

Head, Simon. 2014. *Mindless: Why Smarter Machines Are Making Dumber Humans*. New York: Basic Books.

Hull, Gordon, and Frank Pasquale. 2017. "Toward a Critical Theory of Corporate Wellness. *BioSocieties* 13, no. 1: 190–212. https://link.springer.com/article/10.1057%2Fs41292-017-0064-1.

Innovating Communication in Scholarship. 2015. *Gaming Metrics: Innovation and Surveillance in Academic Misconduct*. December 11, 2015. http://icis.ucdavis.edu/?p=826.

Khurana, Rakesh. 2010. *From Higher Aims to Hired Hands: The Social Transformation of American Business Schools and the Unfulfilled Promise of Management as a Profession*. Princeton, NJ: Princeton University Press.

———. 2016. "Management Theory Is Becoming a Compendium of Dead Ideas." *Economist*, December 17, 2016.

King, Andrew A., and Baljir Baatartogtokh. 2015. "How Useful Is the Theory of Disruptive Innovation?" *MIT Sloan Management Review*, September 15, 2015. http://sloanreview.mit.edu/article/how-useful-is-the-theory-of-disruptive-innovation/.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.

Kleinke, J. D. 2009. "Perfection in PowerPoint." *Health Affairs* 28, no. 4: 1223–24.

Knapton, Sarah. 2016. "One in Three Heart Surgeons Refuse Difficult Operations to Avoid Poor Mortality Ratings, Survey Shows." *Telegraph*, June 3, 2016. http://www.telegraph.co.uk/science/2016/06/03/one-in-three-heart-surgeons-refuse-difficult-operations-to-avoid/?WT.mc_id=tmg_share_tw.

Kronman, Anthony. 1993. *The Lost Lawyer: Failing Ideals of the Legal Profession*. Cambridge, MA: Belknap Press.

Lambert, Craig. 2014. “Disruptive Genius.” *Harvard Magazine*, July/August 2014. http://harvardmagazine.com/2014/07/disruptive-genius.

Larson, Magali Sarfatti. 1977. *The Rise of Professionalism: A Sociological Analysis*. Berkeley: University of California Press.

Lepore, Jill. 2014. “The Disruption Machine: What the Gospel of Innovation Gets Wrong.” *New Yorker*, June 23, 2014. http://www.newyorker.com/magazine/2014/06/23/the-disruption-machine.

MacDonald, Keith. 1995. *The Sociology of the Professions*. Thousand Oaks, CA: Sage.

Madison, Kristin. 2009. “The Law and Policy of Health Care Quality Reporting.” *Campbell Law Review* 31, no. 2: 215–55.

Markides, Kyriakos S., and Jeannine Coreil. 1986. “The Health of Hispanics in the Southwestern United States: An Epidemiological Paradox.” *Public Health Reports* 101, no. 3: 253–65.

McAfee, Andrew, and Erik Brynjolffson. 2017. *Machine Platform Crowd*. New York: W. W. Norton.

McGarity, Thomas O., and Wendy E. Wagner. 2008. *Bending Science: How Special Interests Corrupt Public Health Research*. Cambridge, MA: Harvard University Press.

Mehlman, Maxwell. 2012. “Professional Power and the Standard of Care in Medicine.” *Arizona State Law Journal* 44: 1165–1236.

Mirowski, Philip. 2011. *Science-Mart: Privatizing American Science*. Cambridge, MA: Harvard University Press.

———. 2014. “The Political Movement that Dared Not Speak Its Own Name: The Neoliberal Thought Collective under Erasure.” Institute for New Economic Thinking Working Paper

Series No. 23, September 2014. http://cms.ineteconomics.org/uploads/papers/WP23-Mirowski.pdf.

Mirowski, Philip, and Edward Nik-Khah. 2017. *The Knowledge We Have Lost in Information: The History of Information in Modern Economics*. New York: Oxford University Press.

Morgan, David. 2016. "How Psychoanalysis Can Help Us Make Sense of Brexit." *British Politics and Policy* (blog). October 26, 2016. http://blogs.lse.ac.uk/politicsandpolicy/how-psychoanalysis-can-help-us-make-sense-of-brexit/.

Muller, Jerry Z. 2018. *The Tyranny of Metrics.* New York: Princeton University Press.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Pasquale, Frank. 2007a. "Technology, Competition, and Values." *Minnesota Journal of Law Science and Technology* 8, no. 2: 607–22.

———. 2007b. "The Three Faces of Retainer Care: Crafting a Tailored Regulatory Response." *Yale Journal of Health Policy, Law, and Ethics* 7, no. 1: 39–97.

———. 2010a. "Cognition-Enhancing Drugs: Can We Say No?" *Bulletin of Science, Technology and Society* 30, no. 1: 9–13.

———. 2010b. "Reputation Regulation: Disclosure and the Challenge of Clandestinely Commensurating Computing." In *The Offensive Internet: Privacy, Speech, and Reputation*, edited by Saul Levmore and Martha C. Nussbaum, 107–23. Cambridge, MA: Harvard University Press.

———. 2011. "The Troubling Consequences of Trade Secret Protection of Search Engine Rankings." In *The Law and Theory of Trade Secrecy: A Handbook of Contemporary*

Research, edited by Rochelle C. Dreyfuss and Katherine J. Strandburg, 381–405. Northampton, MA: Edward Elgar.

———. 2012. "Accountable Care Organizations in the Affordable Care Act." *Seton Hall Law Review* 42: 1371–92.

———. 2015a. "Synergy and Tradition: The Unity of Research, Service, and Teaching in Legal Education." *Journal of the Legal Profession* 40, no. 1: 25–48.

———. 2015b. "The Algorithmic Self." *The Hedgehog Review* 17(1), 30-46.

Pasquale, Frank, and Glyn Cashwell. 2015. "Four Futures of Legal Automation." *UCLA Law Review Discourse* 63: 26–48.

Peck, Don. 2013. "They're Watching You at Work." *Atlantic*, December. https://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/.

Porth, Eric, Kimberley Neutzling, and Jessica Edwards. n.d. "Functionalism." University of Alabama Department of Anthropology. Accessed July 31, 2017. http://anthropology.ua.edu/cultures/cultures.php?culture=Functionalism.

Putnam, Hillary. 1988. *Representation and Reality*. Cambridge, MA: MIT Press.

Reagle, Joseph. 2014. *Reading the Comments: Likers, Haters, and Manipulators at the Bottom of the Web*. Cambridge, MA: MIT Press.

Reid, T. R. 2010. *The Healing of America: A Global Quest for Better, Cheaper, and Fairer Health Care*. New York: Penguin.

Remus, Dana. 2015. "Reconstructing Professionalism." UNC Legal Studies Research Paper No. 2676094. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2676094.

Rutgers Graduate School Faculty. 2016. "Resolution to Be Considered by the Faculty of the Graduate School, Rutgers University–New Brunswick." May 10, 2016. http://rutgersaaup.org/sites/default/files/images/AA-resolution-for-GSNB-10-May-2016.pdf.

Sadowski, Jathan, and Frank Pasquale. 2015. "The Spectrum of Control: Toward a Social Theory of the Smart City." First Monday 7:20.

Sauder, Michael, and Wendy Nelson Espeland. 2009. "The Discipline of Rankings: Tight Coupling and Organizational Change." *American Sociological Review* 74, no. 1: 63–82.

Schauer, Frederick. 1990. "Rules and the Rule-Following Argument." *Canadian Journal of Law and Jurisprudence* 3, no. 2: 187–96.

Styhre, Alexander. 2013. *Professionals Making Judgments: The Professional Skill of Valuing and Assessing*. (London: Palgrave MacMillan).

Taylor, Charles. 1989. "Cross-Purposes: The Liberal-Communitarian Debate." In *Liberalism and the Moral Life*, edited by Nancy L. Rosenblum, 159–82. Cambridge, MA: Harvard University Press.

Tilson, Hugh, et al. 2016. "Recommendations for Selecting Drug-Drug Interactions for Clinical Decision Support." *American Journal of Health-System Pharmacy* 73, no. 8: 576–85.

Vaheesan, Sandeep, and Frank Pasquale. 2018. The Politics of Professionalism. *Annual Review of Law and Social Science*.

Van Loo, Rory. 2016. "The Corporation as Courthouse." *Yale Journal on Regulation* 33, no. 547: 547–602.

Vinsel, Lee. 2015. "Snake Oil for the Innovation Age: Christensen, Forbes, and the Problem with Disruption." *Taming the American Idol* (blog), September 16, 2015.

http://leevinsel.com/blog/2015/9/16/snake-oil-for-the-innovation-age-christensen-forbes-and-the-problem-of-disruption.

Weizenbaum, Joseph. 1976. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: W. H. Freeman.

Wennberg, John E., Elliott S. Fisher, Jonathan S. Skinner, and Kristen K. Bronner. 2007. "Extending the P4P Agenda, Part 2: How Medicare Can Reduce Waste and Improve the Care of the Chronically Ill." *Health Affairs* 26, no. 6: 1575–85.

Wing, Jeannette M. 2006. "Computational Thinking." *Communications of the ACM* 49, no. 3: 33–35.

Woolf, Stephen. 1998. "Do Clinical Practice Guidelines Define Good Medical Care? The Need for Good Science and Disclosure of Uncertainty When Defining 'Best Practices.'" *Chest* 113, no. 3 Supplement: 166S–171S.

BIO

Frank Pasquale is an expert on the law of artificial intelligence, algorithms, and machine learning. He has been recognized as one of the ten most cited scholars in health law in the United States. His book, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Harvard University Press, 2015), develops a social theory of reputation, search, and finance, and offers pragmatic reforms to improve the information economy. In privacy law and surveillance, his work focuses on the regulation of algorithmic ranking, scoring, and sorting systems, including credit and threat scoring.