Journal of Motor Behavior, Vol. 42, No. 6, 2010

RESEARCH ARTICLE

Experimenting with Theoretical Motor Neuroscience

Robert Ajemian[1], Neville Hogan[2,3]
[1]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge. [2]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge. [3]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge.

ABSTRACT. Motor neuroscience is well over 100 years old, with seminal work such as G. T. Fritz and E. Hitzig's discovery of motor cortex occurring in 1870. Theoretical motor neuroscience has been ongoing for at least the last 50 years. How mature a scientific discipline is motor neuroscience? Are experimentalists and theoreticians working together productively to help the field progress? This article addresses these questions by advancing the following theses. Motor neuroscience remains at a descriptive stage due to the incredible complexity of the problem to be solved. The proliferation of models—and distinct modeling camps—stems from the absence of unifying conceptual constructs. To advance the field, theoreticians must rely more heavily on the concept of falsification by producing models that lend themselves to clear experimental testing.

Keywords: Equilibrium Point Theory, falsification, Optimal Feedback Control, theoretical motor neuroscience

Thirty years ago, theoretical motor neuroscience did not exist as an independent discipline. Since then, the number of theoretical motor neuroscientists has rapidly increased—and keeps increasing—to the point where multiple satellite conferences involving hundreds of individuals (including the one which prompted this special issue) are dedicated to the topic each year. What kind of progress is theoretical motor neuroscience making? Is the field in its infancy or can it be considered mature? We attempt to address these fundamental questions directly for a general neuroscience audience. At the same time, we acknowledge the difficulty of assessing progress in one scientific field without comparing it to progress that has occurred in other scientific fields. For that reason, we begin the article with a general discussion of the stages of scientific progress and the role of theory in mediating that progress, as illustrated by specific historical examples. Using these examples to formulate a general notion of scientific progress, we conclude with our personal evaluation of theoretical motor neuroscience.

Stages in Scientific Progress

All scientists share a common goal: to make sense of the world in which we live. In any domain of inquiry—from the subatomic to the behavioral—we are confronted with a staggering array of observations that overwhelm the untrained eye. It is the job of the scientist to tame this confusion by uncovering rules or principles that codify large amounts of data within compact explanatory frameworks. These principles are referred to as models or theories, and they render the world around us more comprehensible than it would be otherwise.

Not all models are created equal; they vary in their explanatory scope. Some models amount to little more than phenomenological narratives, whereas others embody abstract theoretical unifications. Is one type of theory better than another? Are there stages of theory formation?

Consider the *Periodic Table of the Elements*, one of the crowning achievements in Chemistry. Initially, chemists such as Dmitri Mendeleev (1869) grouped elements according to the peculiar structure of the periodic table because when they did so, regularities appeared in the observed properties of elements—that is, elements arranged in vertical groupings exhibited similar properties such as solubility, reactivity, melting point, and the like. The table was, therefore, initially conceived as a useful heuristic that organized a broad array of observations. Roughly 50 years after the *Periodic Table* came into being, Erwin Schrodinger's famous equation was proposed as a means to explain the behavior of matter at the atomic scale (Schrodinger, 1926). This equation explained the gross structure of the *Periodic Table* as well as much of its fine structure, not to mention a host of nonchemical phenomena involving the behavior of matter under the influence of forces—after all, the equation was designed for this far more general purpose. Therefore, our understanding of the periodic table became demonstratably deeper with the development of quantum chemistry—more observations could be explained with fewer principles.

Everything should be made as simple as possible, but not simpler.

—Albert E. Einstein

In general, scientists want their theories to explain as much as possible as succinctly as possible. The extent to which a given scientific discipline is characterized by a concise set of abstract principles, as opposed to a wide array of phenomenological heuristics, can be taken as a measure of the field's maturity. One could even say that the dream of all scientific disciplines is a grand unifying theory: a concise set of principles or equations from which all observations can be subsequently derived. Obviously, no one knows whether such grand unifying theories are attainable, but this platonic ideal does enable us to establish a developmental axis along which scientific theories progress from simple heuristics toward unifying principles.

The authors contributed equally to this article.

Correspondence address: Robert Ajemian, McGovern Institute for Brain Research, Building 46, Room 6193, 77 Massachusetts Ave., Cambridge, MA 02139, USA. e-mail: ajemian@mit.edu

With this goal in mind, we classify scientific progress broadly in three stages: *heuristic description*, *curve fitting*, and *theoretical synthesis*. During the heuristic description stage, investigators accumulate data in the form of observations that broadly encompass the scope or boundaries of their disciplinary interest, while efforts are made to organize those data in terms of convenient heuristics. Later, as scientists become more familiar with the structure of their data, a curve-fitting stage arises in which portions of the data are modeled with broadly useful formalisms, often mathematically rigorous. Finally, as a field matures, there emerges theoretical synthesis: the development of an elegant and concise set of abstract, yet interpretable, principles—often highly mathematical in nature—from which a large portion of known observations can be derived.

The language of mathematics reveals itself unreasonably effective in the natural sciences . . . a wonderful gift, which we neither understand nor deserve.

—Eugene P. Wigner

We acknowledge that this classification scheme is somewhat arbitrary and that many other schemes would suffice equally well. Moreover, even though mathematics has proven to be a remarkably efficacious tool for understanding the physical world, it is by no means necessary that unifying principles must be deeply rooted in mathematics. Some of the most profound unifications in science have been rendered qualitatively. The best example is Darwin's Theory of Evolution (Darwin, 1859), a succinct—even terse—set of nonmathematical principles that codifies and unifies a tremendous wealth of observations across a multitude of disciplines. These qualifications aside, we reiterate our more fundamental assertion, which is that sciences begin with a descriptive, taxonomic phase and progress naturally by moving toward grand unifying theories. The history of celestial mechanics illustrates this point.

Celestial Mechanics: A Case Study

Heuristic Description

While there were earlier contributions, the foundations of modern celestial mechanics were laid by Tycho Brahe, a Danish nobleman who studied at the University of Copenhagen and developed a precocious and lasting interest in astronomy. Keenly aware of the value of observation, he developed instruments of unprecedented accuracy and built four observatories (at Herrevad Abbey, Ljungbyhed; Uranienborg and Stjerneborg on Hven; and at Benátky nad Jizerou, near Prague) during the course of his lifetime (see, e.g., Dreyer, 1890). Though an avowed empiricist, Brahe was not averse to theorizing, and in the late 16th century he published the Tychonic system, a hybrid geocentric and heliocentric model of the solar system (Blair, 1990). Nonetheless, his fame and lasting impact derived from his remarkable compilation of meticulous stellar and planetary observations, which made an essential contribution to the establishment of modern empirical science.

There are just as many measurements and methods as there are astronomers, and all of them disagree.

—Tycho Brahe

Among Brahe's seminal contributions to astronomy, he was the first astronomer to regularly calibrate his instruments, he was the first astronomer to correct for atmospheric refraction, he was the first astronomer to include estimated error in his measurements, he measured the positions of the planets frequently and at regular intervals enabling detailed orbital reconstructions, he made exhaustive observations of a Supernova (in 1572) occurring within the constellation Cassiopeia, and he determined that the flight path of comets was not situated in Earth's atmosphere (as had been believed) and that comets were actually located beyond the moon. For our purposes, Brahe's most invaluable contributions are found in his scrupulously detailed orbital reconstructions of the planets, data that Kepler put to good use.

Curve Fitting

In his last years, Tycho Brahe hired an assistant, Johannes Kepler, to help with mathematical calculation. Kepler was already an established academic who (since 1594) had taught mathematics and astronomy at the precursor to the University of Graz. With access to Tycho Brahe's prolific observations, Kepler struggled for almost a decade to describe the orbital motion of Mars. He attempted to describe variations in the speed of Mars along its orbit using the Ptolemaic concept of equants.[1] Initially, he succeeded in describing the available observations within average experimental error, which Tycho Brahe had taken pains to quantify. However, Kepler went further to test whether this model correctly predicted observations of the distance from Mars to the Sun (these data had not been used to fit the model). The model predictions differed from observations by more than 40%, well beyond experimental error.

Kepler's solution to this problem ultimately led to his famous three laws of planetary motion (Kepler, 1609, 1619): (a) planets move around the sun in ellipses, with the sun at one focus; (b) the line connecting the sun to a planet sweeps equal areas in equal times; and (c) the square of the orbital period of a planet is proportional to the cube of its mean distance from the sun. These laws constitute a startlingly simple yet spetacularly effective description of planetary motion. On the other hand, they seem disjoint and unrelated. It is hard to see what one law has to do with the others—surely some deeper, causal explanation exists. Kepler himself tried to derive these laws from the idea that the sun, acting as a rotating magnetic monopole, somehow pushes the planets in their orbits; this idea did not succeed.

In modern parlance, Kepler's work could be considered curve fitting because he found three distinct regularities in the planetary data, each of which he described with simple, though somewhat ad hoc, mathematical models. While the term *curve fitting* carries with it the negative connotation of being unprincipled, curve fitting nonetheless embodies a highly meaningful stage of scientific progress, whereby complex data sets can be rendered comprehensible in a flash of insight that seems to proceed ex nihilo. In the proper scientific context, a clever curve fit constitutes a genuine scientific breakthrough, as Kepler's three laws demonstrate to this day. Of course, an excessive reliance on curve fits to explain data can be counterproductive, particularly if the curve fit only holds for a small (or otherwise restricted) data set.

Theoretical Synthesis

In one of the greatest leaps of insight in human history, Isaac Newton formulated a unified theory from which Kepler's curve fits could be derived. This theory, published in the *Principia*, consisted of a law of universal gravitation together with three general laws of motion. From these principles, Newton (1687) elegantly proved Kepler's laws—and much, much more. Newton showed that the laws describing the motion of heavenly bodies also describe the motion of objects on Earth, a profound and spectacular generalization and the reason for the term *universal gravitation*. This unification of celestial and terrestrial mechanics was subsumed under the rubric of classical mechanics and is still a fundamental building block of modern engineering and scientific curricula.

The scope of Newton's synthesis vastly surpassed all that came before. The Heavens and Earth no longer stood as separate realms of hierarchical importance, whose structure was preconceived through a tradition of ancient myths. Rather, all that could be seen to exist assumed its rightful role in an egalitarian clockwork universe. The motion of Mars and the motion of a pendulum could accurately be described by the same concepts and the same equations. In the playful words of a British satirist:

Nature and Nature's laws lay hid in night; God said Let Newton be! And all was light.

—Alexander Pope

Falsification as the Key to Progress

How is scientific progress made? How are transitions made between the various stages of a scientific discipline? Given that scientists want to explain as much as possible as succinctly as possible, it makes sense that science progresses when a new theory promises to explain more than a previous theory in a more succinct manner. Kepler was able to encapsulate Brahe's data within three simple laws. But in so doing, Kepler did not merely repackage Brahe's data; he also generalized Brahe's data by making predictions as to what new data on the motion of planets should look like when astronomers made new measurements. Confirmation of those predictions would provide additional credence for the laws, while disproof of those predictions would falsify the model and encourage a search for a new (or at least revised) model.

A similar story applies to the transition from Kepler to Newton. While the derivation of Kepler's three laws from Newtonian mechanics marks a true epiphany in the history of human thought, Newton's laws apply to far more than planetary motion. Newtonian mechanics makes predictions regarding virtually every terrestrial and celestial physical phenomenon. Yet we know today that Newtonian mechanics does not accurately describe reality at all spatiotemporal scales—the theory of general relativity supersedes it when speeds are very high and masses very large, while quantum field theory supersedes it at the nanoscale and smaller. Nonetheless, while 20th-century physics has falsified Newton's laws, Newtonian mechanics has been rendered neither obsolete nor valueless; rather, it has been refined by delineating its domain of competence, a domain that even today encompasses the vast range of human experience. Planes, trains and automobiles are all designed using Newtonian mechanics. In science, successful theories are not entirely replaced but, rather, incorporated as part of the foundation of newer theories.

Thus, a hallmark of theoretical progress in science is broadening the domain of description, clarifying the domain to which predecessor theories apply and thereby making novel predictions for comparison with new data. These predictions can either be confirmed, adding credence to the theory, or falsified, encouraging a search for a new or revised theory. Falsification, then, is the engine that drives scientific progress by requiring that each new theoretical innovation must extend the scope of scientific explanation to a novel domain of experimental inquiry. Karl Popper is the philosopher who first formally articulated the importance of falsification in science (Popper, 1959).

Every genuine test of a theory is an attempt to falsify it, or to refute it. Testability is falsifiability.

—Karl R. Popper

Without such testing, progress cannot be made. If the enterprise of science is viewed as an attempt to induce Nature to reveal her secrets, then theories are the means by which scientists pose yes/no questions to Nature; experiments designed to falsify theories are the means by which Nature is compelled to answer those questions "yes" or "no."

In a certain sense, the value of a scientific theory depends more on the cleverness of the questions it poses than on the number of affirmative answers it receives. This point is illustrated by another pillar of classical physics, the theory of electromagnetism put forth by James Clerk Maxwell. Prior to Maxwell, numerous investigators—including Ampère, Weber, Gauss, Biot, Savart, Lenz, Faraday, Henry, and Green, among others—had studied various aspects of electricity and magnetism. Building on these heuristic descriptions and curve fits, Maxwell postulated a grand synthesis of electricity and magnetism into a unitary entity of electromagnetism, succinctly described with field equations.[2] Maxwell's (1865) electromagnetic theory extended well beyond the domain of existing descriptions of electric and magnetic phenomena. In particular, it made two bold predictions, one of which was spectacularly verified; the other was falsified in an equally spectacular fashion.

Maxwell's theory predicted that light, the physical phenomenon at the center of the study of optics, was actually nothing more than a traveling electromagnetic wave. Although this idea is taken for granted today, it was a radical departure from contemporary thinking when Faraday and Maxwell proposed it. In 1887, Heinrich Hertz (1893) confirmed this prediction with a series of experiments in which radio waves were generated by electrical circuits.[3] Thus, with one elegant prediction, verified by a series of hypothesis-driven experiments, Maxwell's theory led to the unification of optics, electricity, and magnetism, three fundamental areas of inquiry in the natural sciences that had been considered separate and distinct phenomena for centuries, if not millennia, prior. The explanatory power of this theory was so grand, that, as the 19th century drew to a close, many prominent thinkers felt compelled to declare an end to the physical sciences—only filling in the details remained.

There is nothing new to be discovered in physics now; All that remains is more and more precise measurement.

—William Thomson (Lord Kelvin), 1900

A second prediction made by the theory was the existence of the ether. Maxwell's equations treat light as an electromagnetic wave. Its propagation speed, determined by two measureable physical properties (permeability and permittivity), is constant, independent of the motion of its source or of an observer. This prompted the postulate of a luminiferous ether—in essence, the medium that did the waving—which provided a unique, absolute rest frame of reference. That, in turn, implied that the speed of light should appear to vary with the speed of the observer through that medium. In the late 19th century, a series of superbly crafted experiments by Albert Michelson and Edward Morley at Case Western Reserve University provided strong evidence that the apparent speed of light did not vary, as was predicted by ether theories (Michelson & Morley, 1887). Though debate continued for decades,[4] this experimental observation falsified Maxwell's electromagnetic theory. This was perhaps the most influential "failure" in the history of science, ultimately paving the way for development of special relativity by Albert Einstein in 1905.

These historical examples may help to clarify the value of falsification as a part of theory building. It is unfortunate that the term *falsify* has a negative connotation; in reality the outcome of an attempt to falsify is always positive, for it either assures the rectitude of the current path or it closes blind alleys, thereby opening new avenues of inquiry. In either case, science collectively makes progress when nature is being asked insightfully tough questions in a timely manner. (To be sure, the answers to those questions may impact the historical reputations of individual scientists, but that is a separate issue unrelated to scientific progress as a whole.)

You imagine that I look back on my life's work with calm satisfaction, but from nearby it looks quite different. There is not a single concept of which I am convinced it will stand firm, and I feel uncertain whether I am in general on the right track.

—Albert E. Einstein

Where Is Motor Neuroscience?

At what stage of progress is motor systems neuroscience? Given the paucity of broadly applicable theories of how the brain works, it seems reasonable to conclude that no branch of systems neuroscience should be considered at the stage of theoretical synthesis or "grand unification." Is it at the heuristic description stage? The curve-fitting stage? Early in the theoretical synthesis stage? Or somewhere in between? Is the field making progress or is it in a "holding pattern?" If the latter, how can progress best be facilitated?

Even though neuroscience is a relatively young field, an enormous amount of data has been accumulated over the last several decades. At the 1972 conference of the International Society for Neuroscience, there were 1,229 attendees; in 2009, there were roughly 31,000 attendees. In 1972, there were 4 dedicated neuroscience journals. Today the ISI Web of Knowledge lists 231 journals in the neurosciences. By these and many other measures, the volume of published neuroscience data has grown exponentially for several decades. Some of this rapid growth may be attributed to advances in technology for probing brain function (e.g., functional magnetic resonance imaging [FMRI], multielectrode recording arrays, patch clamp recording, genetic techniques like the polymerase chain reaction [PCR], the increasing miniaturization of sensors for obtaining kinematic and kinetic data, and many others). But most of it is due to the related fact that the number of working neuroscientists has increased tremendously. Lest we forget, the U.S. Congress, together with the president, formally declared the 1990s as the decade of the brain.

In concert with neuroscience in general, motor neuroscience has grown exponentially over the last several decades. The sheer volume of data that has been accumulated seems to suggest that the field has progressed beyond the heuristic description stage. This speculation is supported by the vast number of theories and models that have been proposed to codify single data sets or subsets of published data—after all, as the amount of data has grown exponentially, so too has the number of theoretical motor neuroscientists whose job it is to explain the data at all levels. A partial list of some of these codifying ideas could include: equilibrium point control, motor programs, the population vector hypothesis, dynamical systems approach theory, motor primitives for synergistic control, Bayesian state estimation, uncontrolled manifold analysis, stochastic Optimal Feedback Control (OFC), and the like—and this list is by no means comprehensive. Some of these ideas concern neural data, some concern behavioral data, some concern both. Some of these ideas are currently in favor, and some have languished. Some require sophisticated mathematics, though others remain largely intuitive and require minimal mathematics. Given the abundance of theories and models that have been developed to explain different domains of motor control, should one conclude that motor neuroscience has progressed well into the curve-fitting stage, if not further?

A wealth of information creates a poverty of attention.

—Herbert A. Simon

We submit that the answer to that question is a definitive *no*. Although there is an abundance of theorists and theories, no consensus has emerged regarding the applicability of any one theory to any particular segment of data, a requirement for advancing past the heuristic description stage. In the case of celestial mechanics, there emerged a clear consensus regarding the utility of Kepler's laws as the best contemporary explanation of data on planetary orbits. To be sure, these laws did not address numerous other celestial phenomena (and ignored terrestrial phenomena in their entirety). Nor was it clear where Kepler's laws came from. But no post-Keplerian and pre-Newtonian astronomer used any other framework for describing the manner in which the planets appeared to move. In the case of today's theoretical motor neuroscience, each theory constitutes a speculative hypothesis, the utility of which depends largely on the point of view of the beholder. At some level, new ideas are met with debate and disagreement in any field, but what is striking in theoretical motor neuroscience is the extent of such dissent. One could say, in line with the earlier Brahe quote, that there are just as many ideas as to how the nervous system executes movements as there are motor neuroscientists. The result is a balkanization of motor neuroscience into competing "schools of thought," tending to promote "*eminence*-based" explanation over "*evidence*-based" science.

One example is the question of how the motor cortex represents movements. This issue has been debated for well over a century, and despite the simplicity of the question, perspectives on the issue are just as thoroughly split today as they were a hundred years ago. In the late 1800s, Fritsch and Hitzig (1870) believed that the motor cortex encoded the muscular details of movements, whereas Ferrier (1873) believed that the motor cortex represented high-level movement commands. Today some researchers still believe that the motor cortex encodes the low-level muscular details of movement commands, while other researchers maintain just as ardently that high-level movement commands, such as hand velocity, are represented (e.g., Ajemian et al., 2008; Georgopoulos, 1995; Loeb, Brown, & Scott, 1996; Scott, 2000; Taylor & Gross, 2003). These debates do not persist for lack of experimental investigation during the intervening time period; indeed, hundreds of papers have been published addressing precisely this question. Rather, our state of knowledge is such that different individuals may interpret the same data set completely differently.

Because of this ongoing and omnipresent lack of consensus, theoretical motor neuroscience (and systems neuroscience more generally) rapidly has devolved, shortly after its birth, into a feudal patchwork of independent fiefdoms, each espousing and advocating its own unique conceptual infrastructure, which, like a medieval crest, serves to unify thematically the efforts of all who toil under its banner. Unfortunately, we have yet to emerge from this inaugural balkanization. Does this reflect poorly on the field? Not necessarily. Just as species in nature have to compete for survival, so too do ideas in science, particularly at the early developmental stages of a scientific discipline. From this perspective, the existence of splinter groups pursuing competing paradigms may reflect the fact that the creative process is alive and well in a field that is currently in the late heuristic description stage.

A scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die and a new generation grows up that is familiar with it.

—Max Planck

That said, should one find it disheartening to work in a field with so little consensus? Should we give up (for now) and try working on other problems? Again, we believe not, this time for two important reasons: (a) understanding how the mind arises from the brain is one of the last great mysteries in all of science, and (b) neurological diseases exact an exorbitant health cost on our population, and only through a better understanding of how the brain works can we hope to mitigate this cost.

Regarding the first point, the lack of maturity we ascribe to theoretical motor neuroscience in no way signifies a lack of skill on the part of current practitioners. Rather, it serves as a continuing testament to the unprecedented difficulty of the problems motor neuroscientists are trying to solve. The brain embodies, arguably, the most complex aggregation of matter and energy ever studied or known to exist in the universe. Neurotransmission depends on biochemical processes that occur at a spatial scale of nanometers (a synaptic cleft is about 20–40 nanometers wide) and a temporal scale of milliseconds or less. Human motor performance takes place at scales of meters and seconds or greater. Thus, the actualization of purposive human behavior emerges from a highly coupled dynamical system operating over about 12 orders of magnitude across both space and time, and possibly more. Experimental data are gathered at all of these scales, from the molecules involved in synapse formation to the global patterns of brain activity to the motor behavior observed during task performance. To confront the enormity of the resulting data set and the pace at which it expands is daunting—to say the least.

Regarding the second point, a variety of diseases of the motor system, such as Cerebral Vascular Accident (stroke), Cerebral Palsy, Parkinson's Disease, Dystonia, and Amyotrophic Lateral Sclerosis (among many others), exact an enormous human health cost.[5] Despite considerable research into their etiology, these diseases remain poorly understood. The process of recovery following neurological injury is even less understood. Further research in theoretical motor neuroscience is critical to improving our methods of treatment and ameliorating (and potentially curing) these and related diseases. A cascade of translational benefits will result from significant progress in theoretical motor neuroscience, and as a society, we have an interest and obligation to pursue this critical knowledge gap.

How Can Motor Neuroscience Get Where It Needs to Go?

If the difficulty of the problems to be solved places the field late in a heuristic description phase, what contribution may be expected from the hundreds of theoretical neuroscientists currently at work? The answer to this question has not changed since the time Francis Bacon gave birth to the modern version of the scientific method:

- Develop a theory inspired by existing data.
- Work out experimentally testable predictions of this theory.
- When the results of future experiments are available, update the theory accordingly. If the data falsify the theory, either modify it or abandon it.
- Repeat *ad infinitum*.

All science must be testable. To the extent that theory contributes to science, its contribution must similarly be testable. Falsification provides the ultimate test, even when definitively controlled experiments are problematical or even impossible. Consider paleontology, a field in which experimentation is challenging in the extreme. The theory of evolution constitutes one of its unifying principles, and, though evolution continues to be debated to this day, it makes unambiguously falsifiable predictions. When challenged to identify a critical test (falsification) of the theory of evolution, British biologist Jack (J. B. S.) Haldane famously retorted "Fossil rabbits in the precambrian." In a field like motor neuroscience that has yet to develop a sound theoretical foundation, falsification should, more than ever, serve as a guiding principle for attempts at theory building.

Not all theoretical motor neuroscientists agree with this perspective, and several speakers at the conference raised objections to falsification being the guiding principle of theory development. While only those individuals themselves can properly represent their arguments—and video of the entire conference is online at the National Institute of Health's NINDS Web site—the basic line of thought went something like this:

> Yes, falsification is useful in physics. But physics is a clean scientific discipline with clearly defined concepts that can be instantiated through tangible measurements. Neuroscience, however, is a more messy and complicated field. Therefore, it is unrealistic to evaluate theories in neuroscience with a criterion of falsification, because none will survive. What we should be aiming for, rather, are theories that somehow or another do A LOT—i.e., explain a lot of data with few parameters, provide explanatory frameworks for the brain's remarkable competencies, etc.—without worrying in strict terms about whether the theory generates any testable predictions.

Clearly, this argument has some resonance. Neuroscience is not like physics. Systems neuroscience presents unique challenges, having to bridge an unprecedented range of scales and complexities. Nonetheless, the complexity of systems neuroscience, even if unparalleled, does not, by itself, constitute a legitimate basis for taking liberties with the scientific method. After all, the scientific method has served Western science extremely well during the last several centuries of naturalistic inquiry, and to change the rules of the game now, after encountering some unexpectedly hard problems, seems epically shortsighted for many reasons, of which we discuss two.

> *This technique, of soliciting many modest contributions to the store of human knowledge, has been the secret of Western science since the seventeenth century, for it achieves a corporate, collective power that is far greater than one individual can exert.*
>
> —John M. Ziman

First, as illustrated previously, falsification provides an epistemological impetus for science to progress. To abandon falsification as a guiding principle would be to abandon the goal of putting systems neuroscience on equal footing with other more advanced scientific disciplines. Of course, embracing the criterion of falsification is no guarantee that systems neuroscience will ever be a solved problem. After all, those who advocate a de-emphasis on falsification may ultimately be right about the inherent intractability of the problems we face: it may never be possible to understand the brain the same way we understand other physical phenomena, such as atoms, chemical reactions, or DNA. Nevertheless, adhering to the scientific method is our best chance for ensuring reliable progress, and our belief is that we should spend the next few centuries (or whatever it requires) trying to put neuroscience on par with physics using this time-tested approach. If, after that time, the scientific method has still not yielded the desired conceptual progress in systems neuroscience, then perhaps alternative methodological approaches will have emerged and should be considered.

Second, an essential attribute of science is its objective character. Certain statements about nature are not open to debate or interpretation because their veracity can be demonstrated (at least within the limitations of available observations). For scientific theories to maintain an objective character, they must be validated by the truth or falsity of the predictions they make. If a theory is not testable, then the knowledge it seeks to embody must be removed from any discourse about objective reality because no such designation could be made one way or another. Without recourse to objective testability, the relative merit of competing scientific theories could never be characterized objectively. Rather, the better theory would be determined by caprice, such as current popularity, trendiness, the political cache of its practitioners and the like. Just as certain works of art ebb and flow in reputation over the centuries, so too would the status of competing theories—and that would belie the essence of scientific progress, whereby gains in our understanding of nature are consolidated and accumulated from one generation to the next.

At the end of the day, we are left with the fact that motor neuroscience specifically—and systems neuroscience generally—is really, really hard. At present, we see no way around this fact and no meaningful shortcuts to take. Is there anything in this assessment that can help guide future efforts? We argue that the difficulty of motor neuroscience should compel the motor neuroscientist to return to epistemological basics, as befitting a scientific discipline in the late heuristic description stage: look to develop new terms, new distinctions, new methods of data collection, and, most importantly, new applications of mathematics. Basically, scientists need to do whatever it takes to try to make more sense of the data than has been made heretofore—as long as that route leads to testable predictions.

Never give in—never, never, never, never.

—Winston Churchill, 1941

From one perspective, returning to basics can be frustrating and discouraging, a seeming admission of an overall lack of progress (particularly as other scientific fields, such as molecular biology, appear to make breakthroughs routinely). But from another perspective, working in a late heuristic description phase can be exciting and invigorating. It naturally injects more creativity ("art," if you will) into the process of science, for the scientist is no longer confined to solving puzzles within the tight constraints of a dominant paradigm. Instead, considerable room for innovation exists when the conceptual elements of a scientific field are open to reformulation.

Of course, to be resolutely pragmatic, efforts to lay the conceptual groundwork of a discipline may not necessarily further one's career; an extraordinary amount of time and energy would be required with no guarantee of an eventual payoff. Funding agencies rarely operate by this model, nor is academia's "publish or perish" environment conducive to quixotic attempts to restructure a scientific discipline. Moreover, even if one does succeed in devising a new and potentially fruitful paradigm, the chances are good that it will be opposed by the scientific establishment, which has a variety of vested interests in preserving the status quo (Kuhn, 1962). Given this state of affairs, one might argue that in a late heuristic description stage, incentives exist, ironically, to publish models that *cannot* be falsified, since they can, by definition, be continually recycled in publications without fear of experimental contradiction. But any serious attempt to contribute to scientific progress would not succumb to this temptation.

What I am saying is that, in numerous areas that we call science, we have come to like our habitual ways, and our studies that can be continued indefinitely. We measure, we define, we compute, we analyze, but we do not exclude. And this is not the way to use our minds most effectively or to make the fastest progress in solving scientific questions.

—J. R. Platt

Theory and experiment are methodologically interconnected in the process of scientific discovery: theory should inform experiments and experiments should revise theories. These complementary objectives are most effectively integrated through the method of "strong inference" (Platt, 1964) whereby: (a) possible variants of a theory are explicitly enumerated, and (b) these different alternatives are evaluated through experiments designed to discriminate between them as cleanly as possible. To illustrate this point on how motor neuroscience can best make progress, we use the next section to compare two widely discussed theories of motor control, Equilibrium Point Theory (EPT) and OFC. From our perspective, these theories appear to differ in the extent to which the notion of falsifiability—both in practice and in spirit—is used to guide reciprocal (and beneficial) interaction between theory and experiment.

EPT versus OFC

The core idea behind EPT is simple, yet powerful: movement emerges from posture. With this assumption, the messy computations necessary to solve the inverse dynamics problem disappear. The body knows, presumably, how to program postures, and movement can be represented as a temporal sequence of postures blended together. Generally speaking, there are two versions of EPT, the lambda version (attributable to Feldman, 1966, 1986) and the alpha version (attributable to Bizzi et al., 1982, 1984). These models differ in the extent to which spinal mechanisms (lambda) versus peripheral mechanisms (alpha) are thought to contribute to the attainment of postural targets. Nonetheless, in both cases the nervous system, capitalizing on the known physiological properties of the motor system, centrally specifies a much simpler command than it would be required to specify if all of the joint torques needed to be computed for any given movement.

EPT, like all successful theories, makes testable predictions to guide further research, and many of these experiments have been performed. The basic paradigm for testing EPT involves perturbing the motor system through a variety of means (e.g., modifying proprioception, altering movement dynamics) and observing whether or not the motor system can recover the endpoint goal. If movement is planned as a temporal sequence of postures, then perturbations in the middle of the trajectory should not prevent the subject from ultimately reaching the goal. Indeed, under many circumstances, subjects do indeed still reach the endpoint goal despite significant perturbations (e.g., Bizzi et al., 1984).

However, recovery of the endpoint goal is not possible under all circumstances (e.g., Lackner & Dizio, 1994), and while variants of EPT still effectively explain a large amount of data (McIntyre & Bizzi, 1993), it seems clear that EPT is not the final answer to the question of movement control (though see Feldman & Latash, 2005). Situations exist where the dynamics of movements are, for the most part, preplanned. This said, EPT has left a lasting legacy on the field of motor control for a variety of reasons. First, it embodies the first serious effort to reconcile the computational requirements of movement with the physiological properties of the motor system. In this sense, it is a uniquely biological theory of motor control that has nonetheless been extended as a useful strategy in robotics (Hogan, 1980, 1985; Hogan & Buerger, 2004). Secondly, the theory accurately accounts for behavior in many contexts. For example, when a beginner is learning a sport, it is known that a predominantly positional control strategy is adopted through co-contraction (Wong et al., 2009). When more expertise is acquired, movements develop greater fluidity. Finally, EPT developed the concepts and the vocabulary necessary to tackle some of the fundamental problems of motor control.

EPT gave the field of motor neuroscience quantitative structure and a foothold from which to attempt to solve some very difficult problems. By "bootstrapping" from this foothold, the field made great progress through a variety of experiments designed to test EPT. Like all models, EPT seems not to be a final or complete theory, but a better barometer of a model's import is the number and subtlety of the questions it enabled and answered through experiments. By this measure, EPT must be considered a success.

More recently, there has been a resurgence of interest in using principles of stochastic OFC as a framework for understanding how the brain controls movement (e.g., Todorov & Jordan, 2002). The mathematical foundations of OFC were fully elaborated about fifty years ago by Bellman and Dreyfus (1962), Pontryagin (1962), and others, though numerical procedures for efficient approximate solution continue to be refined to this day. Conceptually, the basic idea is simple: a physical system can be described by some dynamical equations that govern its transitions from one state to the next. The system can be steered by a controller through inputs that affect the state transitions. The controller is tasked with applying inputs to minimize an explicit cost function that depends on the inputs and perhaps the system states (e.g., penalizing deviation from a desired state), using feedback information it receives about the current system state, and often contending with stochastic variation of the control inputs and/or feedback information. For example, when the theory is applied to upper-extremity actions, the arm (and hand) is the physical system to be controlled, the inputs to the system are signals sent from the central nervous system to the muscles, and feedback originates from vision, proprioception, or other afferents. The cost function codifies the task goal (e.g., a desired trajectory or posture of the arm) and typically also penalizes muscle activations.

One of the strengths of the theory is its tremendous generality. OFC was specifically designed by mathematicians and engineers to deal with a large class of problems. It is hardly surprising that the framework can be used to describe locomotion (Nubar & Contini, 1961; Hatze, 1976; Crowninshield & Brand, 1981), posture (Hogan, 1984a), arm movement (Hogan, 1984b) or just about any movement problem. At the conceptual level, several known properties of the motor system can be explained from the framework of OFC (Diedrichsen et al., 2010). For example, the motor system is known to exhibit "structured variability"—that is, when the same goal-directed task is repeated over and over again, significant variability occurs in the kinematics of the trajectory. This observed variability is not random but rather exists in certain patterns. According to OFC, the variability is naturally structured by task-dependent feedback control—deviations relevant to the task goal must be corrected, while deviations irrelevant to the task goal are allowed to accumulate. (Of course, it should be noted that other models provide alternative explanations for structured variability, e.g., Scholz & Schoner, 1999, Cohen & Sternad, 2009).

Ironically, one of the main weaknesses of the theory is also its great generality. The machinery of OFC is so powerful from a mathematical point of view that merely by tweaking one of its many parameters, one can force the model to fit almost any conceivable data set. The framework, therefore, is grossly under-constrained. Even some of the staunchest advocates of OFC have pointed out the extent of this weakness. As Shadmehr and Krakauer (2008) note,

> These results highlight a number of important problems with our framework. First, without knowing precisely the costs and rewards of a movement, it will not be possible to make quantitatively reliable predictions of behavior. Without a priori predictions, how can the theory be falsified? That is to say, if we have experimental results and are allowed to tweak the costs or their weightings until we get a good fit then what have we learned? We would suggest that the best way to proceed is to either specify the cost function before experiments are conducted and make predictions, or fit the costs to data to find the best parameter fit.

We would agree that either the cost function should be specified before the experiment or the cost functions should be determined from data in one experiment and applied to data in another experiment. Unfortunately, we know of no examples where this procedure has been carried out. Worse yet, when those who practice OFC are approached with this criticism, one of their primary responses may be paraphrased as follows: since the brain is highly adaptive, it likely changes strategies from task to task, so the cost functions themselves are "task-dependent" and should not be fixed. This is a classic example of Joseph Heller's famous "Catch-22": the model cannot be tested without fixing the cost function, yet a meta-assertion of the model is that the cost function cannot be fixed.[6]

There is, however, an even deeper problem that holds for the currently fashionable OFC theories. The cost functions include the terms for muscle activations, which embody the system inputs designated as $u_1, u_2 \ldots u_n$. In engineering usage, these inputs represent clearly defined, real quantities that can be measured and whose cost is explicitly known. In the case of OFC applied to motor control, the cost of muscle activation is most often a conceptual placeholder used to cast the problem of movement in familiar terms—muscle activations are not actually measured in most experiments designed to support OFC. Even if muscle activations were measured (e.g., via EMG), as yet no reliable means has been developed to assign a real cost to an EMG value or to compare the costs of different EMG values across different muscles. Unless or until muscle activation costs can be made more real and quantitatively verifiable, a model that includes them is not sufficiently definite to admit experimental testing.

This weakness of the theory is well understood even by its proponents. Guigon, Baraduc, & Desmurget (2008) wrote, "The proposed model is a computational approach to motor control and variability. It is not a model of physiological mechanisms involved in motor control." This common refrain is heard often from those advocating OFC. Basically, they claim that they are trying to understand the behavior of the motor system and not its physiology. While a computational black-box model may indeed provide insight into how the motor system operates, clearly a mechanistic model with physiologically interpretable elements would provide much greater insight.

None of this is to say that OFC is not the ultimate solution to problems in biological motor control. It may be. But before the merit of the model can assessed, it must first be cast in a falsifiable form. Therefore, in line with the quote from Shadmehr and Krakauer (2008), we urge those working on OFC models to focus their efforts on making the models falsifiable, rather than further extending their domain of operation in a nonfalsifiable manner. To quote Karl Popper again, "A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice."

Conclusion

Scientists will always disagree—as well they should. If scientists spoke in unison, the creative energy necessary for the next scientific breakthrough would be lacking. Furthermore, history automatically assumes the role of final arbiter in scientific debates, thereby leaving scientists free to assume an advocacy role for their favored ideas. Nonetheless, this article makes the case that such advocacy should adhere to certain boundaries. Scientists may differ in their ideas, their interpretation of data, and more, but there should be no disagreement about the underlying epistemology of science itself. That is a solved problem. Critical testing—falsification—is, was, and always will be a key element of scientific progress. Centuries of post-Renaissance naturalistic inquiry leave no room for debate on this matter. We must never forget this fact, even when we blindly grope and stumble through the conceptual morass of an extremely challenging scientific discipline mired in its heuristic description phase. As scientists, we must never succumb to the temptation of building elegant models that can never be tested, for such models are not truly scientific.

NOTES

1. Motion along a circle with uniform angular speed about a point (*punctum aequans*) other than its geometric center.
2. Maxwell's initial 20 equations were elegantly reduced to 4 by Oliver Heaviside in an act of mathematical synthesis.
3. This, in turn, laid the foundation for Guglielmo Marconi's development of wireless telegraphy and global radio communication.
4. Neither Michelson nor Morley considered their negative result to disprove the existence of ether.
5. The American Stroke Association estimates that Americans will pay about $73.7 *billion* in 2010 for stroke-related medical costs and disability.
6. As soon as task-dependent differences are allowed, an infinity of experiments is required for strict falsification.

ACKNOWLEDGMENTS

The authors thank Dagmar Sternad for her helpful comments. The meeting focused on discussing this topic was supported, in part, by NIH grant R13-NS065552. Robert Ajemian was supported, in part, by NSF grant #IIS-0904594 for Collaborative Research in Computational Neuroscience. Neville Hogan was supported, in part, by the Eric P. and Evelyn E. Newman Fund.

REFERENCES

Ajemian, R., Green, A., Bullock, D., Sergio, L., Kalaska, J., & Grossberg, S. (2008). Assessing the function of motor cortex: Single-neuron models of how neural response is modulated by limb biomechanics. *Neuron*, *58*(3), 414–428.

Bellman, R. E. & Dreyfus, S. E. (1962). *Applied dynamic programming*. Santa Monica, CA: RAND.

Bizzi, E., Accornero, N., Chapple, W., Hogan, N. (1982). Arm trajectory formation in monkeys. *Experimental Brain Research*, *46*, 139–143.

Bizzi, E., Accornero, N., Chapple, W., & Hogan, N. (1984). Posture control and trajectory formation during arm movement. *Journal of Neuroscience*, *4*, 2738–2744.

Blair, A. (1990). Tycho Brahe's critique of copernicus and the Copernican system. *Journal of the History of Ideas*, *51*(3), 355–377.

Cohen, R. G., & Sternad D. (2009). Variability in motor learning: Relocating, channeling and reducing noise. *Experimental Brain Research*, *193*, 69–83.

Crowninshield, R., & Brand, R. (1981). A physiologically based criterion of muscle force prediction in locomotion. *Journal of Biomechanics*, *14*(11), 793–801.

Darwin C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, England: Murray.

Diedrichsen J., Shadmehr, R., & Ivry, R. B. (2010). The coordination of movement: Optimal feedback control and beyond. *Trends in Cognitive Science*, *14*, 31–39.

Dreyer, J. L. E. (1890) *Tycho Brahe: A picture of scientific life and work in the seventeenth century*. Edinburgh, Scotland: Adam & Charles Black.

Feldman, A. G. (1966). Functional tuning of the nervous system with control of movement or maintenance of a steady posture II: Controllable parameters of the muscle." *Biophysics*, *11*, 565–578.

Feldman, A. G. (1986). Once more on the equilibrium-point hypothesis (Λ model) for motor control. *Journal of Motor Behavior*, *18*, 17–54.

Feldman, A. G., & Latash, M. L. (2005). Testing hypotheses and the advancement of science: Recent attempts to falsify the equilibrium point hypothesis. *Experimental Brain Research*, *161*, 91–103.

Ferrier, D. (1873). Experimental researches in cerebral physiology and pathology. *West Riding Lunatic Asylum Medical Reports*, *3*, 30–96.

Fritsch, G.T., & Hitzig, E. (1870). On the electrical excitability of the cerebrum. In Von Bonin, G. (Trans.), *Some papers on the cerebral cortex (1960)*. Springfield, IL: Thomas.

Georgopoulos, A. P. (1995). Current issues in directional motor control. *Trends in Neuroscience*, *18*, 506–510.

Guigon, E., Baraduc, P., Desmurget, M. (2008). Computational motor control: Feedback and accuracy. *European Journal of Neuroscience*, *27*(4), 1003–1016.

Hatze, H. (1976). The complete optimization of a human motion. *Mathematical Biosciences*, *28*, 99–135.

Hennig W. (1950). *Grundzüge einer Theorie der Phylogenetischen Systematik*. Berlin, Germany: Deutscher Zentralverlag.

Hertz, H. R. (1893). *Electric waves: Being researches on the propagation of electric action with finite velocity through space* (Trans. David Evans Jones). Ithaca, NY: Cornell University Library.

Hogan, N. (1980). Mechanical impedance control in assistive devices and manipulators. *IEEE Joint Automatic Controls Conference*, *1*, TA-10-B.

Hogan, N. (1984a). Adaptive-control of mechanical impedance by coactivation of antagonist muscles. *IEEE Transactions On Automatic Control*, *29*(8), 681–690.

Hogan, N. (1984b). An organizing principle for a class of voluntary movements. *Journal Of Neuroscience*, *4*(11), 2745–2754.

Hogan, N. (1985). Impedance control: An approach to manipulation. *Journal of Dynamic Systems Measurement and Control*, *107*(1), 1–24.

Hogan, N. & Buerger, S. P. (2004). Impedance and interaction control. In Kurfess, T. R. (Ed.) *Robotics and automation handbook* (pp. 19-1–19-24). New York, NY: CRC Press.

Kepler, J. (1609). *Astronomia Nova [New astronomy]*. Donahue, W. H. (Trans., 1992). Cambridge, England: Cambridge University Press.

Kepler, J. (1619). *Harmonice Mundi [Harmony of the worlds]*. Wallis, C. G. (Trans., 1952). Chicago, IL: Encyclopia Britannica.

Kuhn, T. S. (1962). *The structure of scientific revolutions* (1st ed). Chicago, IL: University of Chicago Press.

Lackner J. R., & Dizio, P. (1994). Rapid adaptation to Coriolis force perturbations of arm trajectory. *Journal of Neurophysiology*, *72*(1), 299–313.

Loeb, G. E., Brown, I. E., & Scott, S. H. (1996). Directional motor control. *Trends in Neuroscience*, *19*(4), 137–138.

Maxwell, J. C. (1865). A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, *155*, 459–512.

McIntyre, J., & Bizzi, E. (1993). Servo hypotheses for the biological control of movement. *Journal of Motor Behavior*, *25*(3), 193–202.

Mendeleev, D. (1869). On the relationship of the properties of the elements to their atomic weights. *Zhurnal Russkoe Fiziko–Khimicheskoe Obshchestvo*, *1*, 60–77.

Michelson, A. A., & Morley, E. W. (1887). On the relative motion of the Earth and the luminiferous ether. *American Journal of Science*, *34*(203), 333–345.

Nubar, Y., & Contini, R. (1961). A minimal principle in biomechanics. *Bulletin of Mathematical Biophysics*, *23*, 377–391.

Newton, I. (1687). *Philosophiæ Naturalis Principia Mathematica*. Koyre, A. & Cohen, I. B. (Eds., 1972) Cambridge, MA: Harvard University Press.

Platt, J. R. (1964, October 16). Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, *146*, 347–353.

Pontryagin, L. S. (1962). The mathematical theory of optimal processes. *Interscience*.

Popper, K. R. (1959). *The logic of scientific discovery*. London, England: Taylor & Francis.

Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London, England: Routledge.

Scott, S.H. (2000). Reply to "one motor cortex, two different views." *Nature Neuroscience*, *3*(10), 964–965.

Scholz, J. P., Schöner, G. (1999) The uncontrolled manifold concept: Identifying control variables for a functional task. *Experimental Brain Research*, *126*(3), 289–306.

Schrödinger, E. (1926). An undulatory theory of the mechanics of atoms and molecules. *Physical Review*, *28*(6), 1049–1070.

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, *185*(3), 359–381.

Taylor, C. S., & Gross, C.S. (2003). Twitches versus movements: A story of motor cortex. *Neuroscientist*, *9*(5), 332–342.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*(11), 1226–1235.

Wong, J., Wilson, E. T., Malfait, N., & Gribble, P. L. (2009). Limb stiffness is modulated with spatial accuracy requirements during movement in the absence of destabilizing forces. *Journal of Neurophysiology*, *101*, 1542–1549.

Ziman, J. M. (1969, October 25). Information, communication, knowledge. *Nature*, *224*, 318–324.

Submitted November 29, 2009
Accepted August 31, 2010