Available online at www.sciencedirect.com
ScienceDirect
Journal of Physiology - Paris 101 (2007) 203–213
Journal of Physiology Paris
www.elsevier.com/locate/jphysparis

The science of research: The principles underlying the discovery of cognitive and other biological mechanisms

Alcino J. Silva *

Department of Neurobiology, UCLA, 695 Charles Young Drive South, Los Angeles, CA 90095-1761, United States
Department of Psychiatry, UCLA, 695 Charles Young Drive South, Los Angeles, CA 90095-1761, United States
Department of Psychology, UCLA, 695 Charles Young Drive South, Los Angeles, CA 90095-1761, United States
Brain Research Institute, UCLA, 695 Charles Young Drive South, Los Angeles, CA 90095-1761, United States

Abstract

Studies of cognitive function include a wide spectrum of disciplines, with very diverse theoretical and practical frameworks. For example, in Behavioral Neuroscience cognitive mechanisms are mostly inferred from loss of function (lesion) experiments while in Cognitive Neuroscience these mechanisms are commonly deduced from brain activation patterns. Although neuroscientists acknowledge the limitations of deriving conclusions using a limited scope of approaches, there are no systematically studied, objective and explicit criteria for what is required to test a given hypothesis of cognitive function. This problem plagues every discipline in science: scientific research lacks objective, systematic studies that validate the principles underlying even its most elemental practices. For example, scientists decide what experiments are best suited to test key ideas in their field, which hypotheses have sufficient supporting evidence and which require further investigation, which studies are important and which are not, based on intuitions derived from experience, implicit principles learned from mentors and colleagues, traditions in their fields, etc. Philosophers have made numerous attempts to articulate and frame the principles that guide research and innovation, but these speculative ideas have remained untested and have had a minimal impact on the work of scientists.

Here, I propose the development of methods for systematically and objectively studying and improving the *modus operandi* of research and development. This effort (the science of scientific research or S2) will benefit all aspects of science, from education of young scientists to research, publishing and funding, since it will provide explicit and systematically tested frameworks for practices in science. To illustrate its goals, I will introduce a hypothesis (the *Convergent Four*) derived from experimental practices common in molecular and cellular biology. This S2 hypothesis proposes that there are at least four fundamentally distinct strategies that scientists can use to test the connection between two phenomena of interest (A and B), and that to establish a compelling connection between A and B it is crucial to develop independently confirmed lines of convergent evidence in each of these four categories. The four categories include *negative alteration* (decrease probability of A or p(A) and determine p(B)), *positive alteration* (increase p(A) and determine p(B)), *non-intervention* (examine whether A precedes B) and *integration* (develop ideas about how to get from A to B and integrate those ideas with other available information about A and B). I will discuss both strategies to test this hypothesis and its implications for studies of cognitive function.

Keywords: Science of research; Cognition; Plasticity; Memory; Learning

1. Introduction

In the last 100 years science has grown at a vertiginous pace and it has changed every aspect of the world around us. In the process, scientific discovery has also been transformed from an activity that attracted a few people to an

---

[*] Address: Department of Neurobiology, UCLA, 695 Charles Young Drive South, Los Angeles, CA 90095-1761, United States. Tel.: +1 310 7946609; fax: +1 310 794 7088.
E-mail address: silvaa@ucla.edu

0928-4257/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jphysparis.2007.11.002

industrial occupation that engages a significant percentage of the population and a sizeable fraction of national resources in developed countries. Academic and private research efforts employ millions of people and attract hundreds of billions of dollars world-wide. The National Science Foundation (NSF) estimated that in 2004 the United States investment in research and development (R&D) was approximately $312.1 billion (compared to $152 billion in 1990), with approximately one third of that figure coming from federal sources. The US R&D share of the gross domestic product (GDP) in 2004 was nearly 3% compared to 1.4% in 1953! This growing commitment to R&D is not unique to the US. Most developed countries currently spend more than 2% of their GDP in R&D (Shackelford, 2006). This growing social and economic investment in science has also been paralleled by an increasing dependency of societies on the products of science, from the communication and transportation means that have had an incalculable effect on intellectual and commercial interchanges among societies, to the perennial technological advances that power the world's economic engines. Thus, science and its products have become deeply woven into the fabric of the world's cultures and economies (Tansey and Stembridge, 2005).

Despite the unprecedented growth of budgets for research and development in both the public and private sectors, there have been very few concerted efforts to systematically study and analyze the scientific process and foment the development of methods and strategies to increase its overall efficiency. Scientists use objective methods to analyze their subjects (physics, chemistry, biology, etc.), but they have hardly ever used the same scientific methods to study how they do science itself. Instead, this type of analysis has been the domain of philosophers of science. However well intentioned and inspired these efforts have been, they lack the objectivity, testing and systematic character associated with scientific research. The unprecedented expansion of science in the last 50 years has also been accompanied by a growing awareness among scientists of significant inefficiencies in science, some of which have been uncovered by analyses of the scientific literature (Cole and Cole, 1972). For example, a query in the ISI Web of knowledge (08/06/2006) with the key word "Cognition" yielded more than 26,262 articles; surprisingly, 13,268 of these articles were cited only 2 times or less! Similarly, a query in the ISI Web of knowledge (08/06/2006) with the key words "learning and memory" yielded more than 22,261 articles, 7492 of which were cited 2 times or less! This suggests that a very large percentage of the literature on cognition has had little or no impact on the subject. Unfortunately, there is evidence that this state of affairs is not unique to cognitive research, and it is widespread in science, as I will illustrate below.

Fortunately, a number of developments described in detail later, have taken place that promise to facilitate considerably the scientific analysis of the principles underlying the discovery process. The results of this analysis could have an enormous effect on how we fund, teach, publish and carry out science, and it may be especially useful in the development of "young" disciplines, such as neuroscience, with developing approaches and emerging frameworks.

2. The emergence of the science of scientific research (S2)

2.1. S2: a definition

I propose a new scientific effort called S2 (meaning the science of science). This new scientific effort will be focused on the investigation and testing of general principles of scientific practice. Its goals and methods, as I describe below, distinguish it from its mother discipline, the philosophy of science, that has also been concerned with the analysis of the history and intellectual frameworks that govern science. There are many obvious similarities, but also important differences between these two disciplines: Although both are concerned with ideas about the structure of science, S2 is focused on the systematic testing of these ideas, just as medicine is interested in the systematic testing of ideas about the treatment of disease. The ultimate goal of S2 is the development of pragmatic, validated general principles for increasing the efficiency of science, just as the ultimate goal in medical research is the understanding and systematic development of pragmatic practices that improve the efficiency of prevention and treatments. By contrast, the philosophy of science has traditionally been based on individual insights, opinions and ideas whose validation does not depend on either systematic testing or the general acceptance of other practitioners (i.e., philosophers). For example, the philosopher John Stuart Mill developed a set of five methods (or canons) to guide the establishment of causal relationships. However, these were intuitive logical principles whose validity has never been formally tested.

2.2. The emergence of scientific disciplines

As a proposed new scientific discipline, S2 should go through the developmental stages that characterize the ontogeny of other disciplines. For example, most disciplines, from medicine to economics, have gone through at least two key developmental stages that fall out of common social and organizational forces. The first stage is dominated by isolated efforts of individual practitioners and their lineage of apprentices. In this initial pre-scientific stage core practices are not systematically tested and evaluated, but are instead guided by tradition and mostly implicit principles. The second is a mature stage, in which systematic evaluation, formalization and sharing of principles and ideas are at the very core of the practice. In this second stage, a concerted effort to extract general principles from the shared experience of experts becomes central to the discipline. Consequently, in this stage students are no longer dependent solely on the implicit or explicit knowledge of a given expert, since there is a shared body of knowledge that introduces them to fundamental principles that are central to the practice of their discipline. For example, medical students learn about general principles of medicine that guide medical practice, analysts use economic principles in evaluating investments, chemists derive new compounds guided by the laws of synthetic chemistry, etc.

Scientists lack validated general principles for how to maximize the probability of discovery and improve the chances that their work will have an impact on general knowledge and the work of other scientists. This is because S2 is still in its pre-scientific stage: general proven principles are missing and training on the procedures of scientific discovery is mostly dependent on implicit practices passed on from individual scientists to their students and colleagues. However, the unprecedented growth of science in the last 100 years, and recent key developments in computer science and informatics, as well as the emergence of comprehensive electronic scientific databases, are providing the fundamental impetus for the emergence of S2 as an organized scientific discipline. The history of science suggests that key developments have served as catalysts in the emergence of scientific disciplines. For example, Lavoisier's "*Traité élémentaire de chimie*" (Lavoisier, 1789) triggered the transformation of Alchemy into Chemistry, Adam Smith's "*The Wealth of Nations*" (Smith, 1776) was a key step in building the field of Economics at the dawn of the industrial revolution, Hippocrates and Galen's Herculean efforts to formalize and write about general principles in medical care were the foundation stones of modern medicine, etc. In each of these examples there are two key attributes that forecast the dawn of the second stage in the history of the respective disciplines: an explicit effort to extract and test general principles that guide the practice of the discipline, and an attempt to disseminate these principles so that others can use and build upon them. Thus, the focus is shifted from individual practitioners and their untested ideas, to large groups of practitioners and an evolving body of shared ideas, and more importantly, tested principles that guide their activities.

2.3. The sciences of science

Scientists learn how to identify important problems, devise research strategies, spot unique scientific opportunities, make research choices, and determine when a problem is solved, mostly from years of experience and from implicit examples and interactions with their mentors and closest colleagues. Although methods to determine the properties of celestial bodies, derive new chemical compounds, and analyze the functional organization of the brain, are subject to constant development and shared analysis in science, the innermost workings of the process by which progress on these matters is made (S2) is not.

Disciplines as diverse as dialectics, heuristics, theology, and epistemology have been referred to as the Science of Sciences by various authors throughout the ages. Nevertheless, these disciplines did not involve the systematic analysis of scientific practices, the derivation and *testing* of key principles of scientific activity with the aim of dramatically improving the overall efficiency and productivity of research and development. Hence, S2 is first and foremost a pragmatic discipline whose key goal is to extract and test principles that guide scientific research and increase the efficiency of science, just as the tested laws of chemistry, economics and medicine increased the efficiency and efficacy of their respective disciplines.

2.4. S2 and scientific creativity

It is important to stress that the derivation of general principles of scientific investigation are not meant to undermine the creativity of individual scientists; certainly, the general principles of medicine, economics, and chemistry, for example, do not undermine the creativity of surgeons, investors and molecular biologists. What would medicine be today without the systematic formal efforts of Hippocrates, Galen and the many thousands of others that followed in their steps? Would world economies have ever grown to their current size without the economic systems that we have engineered and systematically tested? Could modern drug development research exist without the systematization of chemistry?

2.5. Evidence for a need for S2

Is not science successful enough already? What is the evidence that there is a need for S2? Although it may be easy to develop an opinion on this subject, without a systematic study of science and its practices it is nearly impossible to objectively answer these questions. For example, there is a growing consensus that there are troublesome inefficiencies in science. Even a cursory look at databases of scientific publications demonstrates that a significant percentage of published articles are either never cited or cited only a very small number of times (<3 times; see for example, Hamilton, 1990, 1991). In the biological sciences alone this may amount to more than 7 million publications that probably took at least 100 billion hours of research time, and involved incalculable physical and economic resources. Even though this vast, but unfortunately, rarely cited effort must have had some influence on scientific knowledge and experimentation, it is difficult to determine what it has added to the growth and development of science. The problem is that without a formal and systematic study of science, without general validated principles that explain the activities of science, without a scientific investigation of the practices of scientists (S2), it is virtually impossible to ever determine what needs improving, whether current practices are optimal, and if not, how things could be made more efficient.

Another hint that understanding scientific practice could improve science is that there is clear evidence of dramatic differences between the productivity of scientists. Although social factors as well as variations in natural abilities such as intelligence, perseverance, creativity, drive, undoubtedly play a role in success, it is almost certain that training on principles of scientific discovery, implicit or not, plays a role in determining success in science. Several informal studies of influential scientists revealed a strong lineage effect suggesting that training has an effect on future success (e.g., Zuckerman, 1967; Kanigel, 1993). The problem is that without a systematic study of the inner workings of science it will be very difficult to ever determine whether Nobel laureate trees are due to social factors (e.g. influential mentors open exclusive doors to their students; brilliance attracts brilliance, etc.), or whether they are due to principles and skills that are passed on from mentor to student and that dramatically increase the probability of discovery. Nevertheless, the power of the mentor–student relation is undeniable: out of the ninety two USA Nobel prize winners in the sciences that had gotten the prize by 1972, 48 of them had worked in junior positions with other Nobel winners (Zuckerman, 1967; Kanigel, 1993)! Is this just the result of the "Matthew effect" (those with a lineage of scientific prominence get an unfair share of exposure and credit in science; Merton, 1968) or the result of practices that are implicitly or explicitly passed on from mentor to apprentice?

2.6. What S2 is and is not

S2 is not about the study of the history or sociology of science (also often referred to, and rightly so, as science of science) or the cognitive processes underlying the work of scientists; these neighboring disciplines will aid and inform S2, but they do not reflect the central goals of S2. The accumulation of heuristics about the practices of scientists, such as those described in George Polya's inspired work "*How to Solve it*" (Polya, 1945) are also not the goal of S2. Although valuable and insightful, these heuristics are not scientifically derived principles that have been either tested or validated. Thus, S2 is not concerned with accumulating untested insights about science from scientists or other luminaries, but on generating and *testing* hypothesis concerning the general principles that underlie successful scientific practice.

Nevertheless, in its simplest form, one of the key ideas underlying S2 is the need to formally understand what successful scientists do, generate general hypotheses about the principles underlying their success, use tools such as electronic databases and computer modeling, to test these ideas, share emerging findings among scientists that can then build upon them, and ultimately use these findings to shape both practices in science and the education of young scientists. This investigative work needs to involve practicing scientists, since their intuitive grasp of their discipline and accumulated practical insights may guide the early days of this emerging discipline. In formalizing general principles of S2 it will be invaluable to have a sense of what pushes science forward; For example, intuitions derived from practical experience will be valuable guides in the arduous and rigorous process of generating specific hypotheses in S2, developing tools and experimental frameworks to test them. By directly involving practicing scientists in this effort, emerging knowledge in S2 will have immediate repercussions on the education and activity of scientists. Even a small increase in the efficiency of science could have an exponential effect on discovery and innovation. Did Galen ever imagine that his pioneering efforts would one-day result in societies with a life expectancy of nearly 80 years? Did Lavoisier ever imagine that chemistry would feed, clothe, house, transport, cure, and entertain the world?

3. Tools for S2

Central to the genesis and growth of any scientific discipline is the development of tools for the study and objective evaluation of relevant phenomena and ideas. For example, the influence and impact of the work of pioneers of memory research such as Ebbinghaus (1885), and Pavlov (1927) can be traced back not only to their insights, but more importantly, to the experimental frameworks that they introduced to test their ideas and develop their concepts. Experimental tools often determine what ideas and hypotheses can be rigorously evaluated, and therefore what aspects of the discipline are first investigated, tested and built upon. Undoubtedly, the same is true of studies in S2. Next, I will review three areas of current development whose tools could be used for S2 studies.

3.1. Discovery informatics

The term "discovery informatics" was first defined by William W. Agresti in 2003, as follows: "Discovery Informatics is the study and practice of employing the full spectrum of computing and analytical science and technology to the singular pursuit of discovering new information by identifying and validating patterns in data." (Agresti, 2003). A key idea in this field is to design computer programs that sift through databases for related information and use this to make new discoveries that follow from previously published, but unconnected observations. In its most idealized conception, discovery informatics models the way scientists think and uses the power and relentlessness of computers to find unrecognized connections in the published literature. Science often progresses by connecting sets of information that evolved independently (i.e. bridging areas or even fields of research). Given enough time scientists will eventually find all of the important connections between areas and fields of study. However, the sheer vastness of the available literature and the astronomical number of all possible connections is such that computers could make a real contribution with this association process. For example, recently "Iridescent", a discovery informatics program, was used to suggest that the drug chlorpromazine, normally used to treat psychiatric problems, would reduce the progression of cardiac hypertrophy (Wren et al., 2004).

S2 hypotheses could be used to modify or even conceive new discovery informatics algorithms like "Iridescent"; then, the validity and efficacy of these S2 hypotheses would be indirectly gauged by the success of these algorithms. For example, insights on the implicit and explicit processes that scientists use to sift through and organize information in large data sets could be evaluated with discovery informatics tools. For example, scientists use implicit and explicit principles to weight the relative value of particular observations or findings in large data sets. This process is of great importance in sifting through large bodies of data, and in identifying data sets that are most relevant for the task at hand. Again, discovery informatics could be useful in testing ideas and hypotheses about this mostly intuitive process, and in turn it could also conceivably serve as a tool to discover strategies and general principles that scientists could use to improve their ability to identify relevant information in large data searches. Thus, principles and findings that make explicit the bases for these intuitive processes could be used to fine tune and sharpen these processes, and could consequently be incorporated in the education of scientists. Similarly, once explicit, these intuitive principles could be conceivably improved upon, implemented in discovery informatics, and scientists could be trained to use them. Although automation will continue to have an important role in data mining, it is unlikely that in the near future computers will fully replace the instincts and intuitions of scientists!

3.2. *Artificial intelligence*

The discovery process has also been modeled with artificial intelligence tools, and robots have even been designed that autonomously perform several rounds of experimentation without outside interference. The "*Robot Scientist*" represented a particularly spectacular incorporation and synthesis of many of these artificial intelligence principles and ideas because of its ability to autonomously perform real world biological experiments (King et al., 2004). The robot starts with information available in the scientific literature, comes up with a set of hypotheses that account for experimental observations related to that literature, chooses an experimental strategy to distinguish between competing hypotheses, designs and executes the key experiments, analyses the results and then repeats the cycle *without* human interference (Whelan and King, 2004). The first implementation of this process targeted yeast genomics and used yeast growth mutants to determine the function of amino acid synthesis genes based on yeast strains lacking those genes. The yeast strains were grown on liquid and the robot automated the pipetting and handling of solutions, conducted assays, read out the results of yeast growth in microtiter plates and incorporated them in its cycles of experimental design and execution. The software part of the set-up included background information relevant to the experiment, a logical inference program, an abduction hypothesis generation code, a deduction code for experimental design, and a program that integrated the whole set-up. Amazingly, the robot outperformed in both accuracy and cost nine graduate students tested on a simulator of those experiments (King et al., 2004; Whelan and King, 2004). Importantly, the design of the Robot Scientist was based on artificial intelligence ideas about the process of discovering. A key concept used in constructing hypotheses to be tested by the robot was abductive inference, a process loosely related to deductive inference. Nevertheless, unlike deduction, which draws connections between available information and does not go beyond it, abductive inference uses available information to generate or extrapolate new facts that can then be tested experimentally. For example, yeast strains need specific amino acids, such as tryptophan, for the synthesis of proteins essential for growth. Without tryptophan yeast cannot grow or multiply. Abductive inference would suggest that if a yeast strain cannot grow it could be because it cannot make tryptophan. Of course, there could be many other reasons, and the lack of a tryptophan synthetic enzyme is simply a possible hypothesis that does not follow necessarily from available information. Modeling studies in artificial intelligence have suggested that abductive and deductive reasoning are central to discovery, and the results of the robot scientist certainly attest to this possibility. Therefore, the robot scientist may not only have important consequences for the automation of certain aspects of scientific discovery, but it can also be used to test general principles of discovery, and may play a key role in S2.

Efforts have also been taken to start to understand specific factors underlying the cognitive neuroscience of scientific reasoning. For example, Schunn and Dunbar (1996) showed that the mechanisms used to solve one scientific problem could affect the solution of another, although the subjects seemed unaware of the transference process. Although the cognitive neuroscience of complex problem solving, including creativity, will affect S2, the subject of S2 is not the mind of the scientist but its practices, the goal is not to understand the mental processes used in the science, but the macro structure of decisions and how they affect progress. Just as the cognitive neuroscience of decision-making is of consequence to economics, so is the cognitive neuroscience of complex decision-making relevant for S2. Both economics and S2 are affected by the dynamics and characteristics of decision-making, but neither discipline centers its studies on it.

3.3. *Scientific databases*

In the last 20 years a number of organizations have taken advantage of advances in computing and the ever decreasing price of data storage to develop comprehensive searchable electronic indexes of scientific publications, such as MEDLINE and INSPEC. By the end of 2005, MEDLINE referenced more than 15 million articles in the biological sciences from more than 4600 journals dating back to the 1950s, and INSPEC listed nearly 10 million articles in physics, engineering and computer science going back to 1898. Importantly, there have also been significant efforts to use these and other databases to evaluate the success of individual scientists and the impact of their contributions. For example, the Institute for Scientific Information (ISI) has developed a searchable database of citations. With this service it is straightforward to determine the number of citations associated with individual articles, scientists, topics, etc. However crude these measures may be, emerging metrics, such as citation indexes, have been used to estimate the influence of specific publications, individual scientists and even institutions (Cole, 2000; Hirsch, 2005). More importantly, these extensive electronic records of science and its products could also be useful in chronicling the development and impact of concepts and ideas in science, characterizing the stages that define the evolution of influential ideas, extracting and testing the principles that govern these processes. In other words, these and other sources of information on the processes and products of science are starting to be used to systematically study science.

4. The *Convergent Four* hypothesis

4.1. Introduction

To better illustrate the ideas presented thus far, I will describe next an S2 hypothesis that I shall refer to as the *Convergent Four.* This hypothesis is based on the idea that experimentation in science has three major goals: first, to define individual phenomena of interest, second, to establish connections between these phenomena, and third, to develop tools that facilitate these goals. The Convergent Four hypothesis is focused solely on the second goal: it outlines four criteria for establishing connections between phenomena of interest in science (see for example, Elgersma and Silva, 1999) and does not pertain to either defining individual phenomena or developing tools. It proposes that there are at least four fundamentally different lines of evidence involved in connecting two natural phenomena, including *negative alteration*, *positive alteration*, *non-intervention* and *integration*. The hypothesis assumes that a key goal of science is to explain natural phenomena by establishing networks of causal connections between them. Single connections do not sit alone but within a network of other connections, with one connection contributing to others in the network. Major theories in science provide explanatory frameworks for these networks of connections, and fields in science define areas of the problem space with clusters of these networks.

It is important to note that this is not the first time that specific criteria are proposed for making connections between natural phenomena. Previous attempts, however, were mostly focused on specific problems such as the case of Koch's Postulates described below (Koch, 1884), which dealt with the relation between pathogens and infectious disease, Rose's requirements for a biochemical mechanism of memory (Rose, 1981), which are criteria for assessing the connection between synaptic plasticity and memory (Martin et al., 2000), etc.

To illustrate the *Convergent Four* hypothesis, and the interdependency of scientific connections, I will use two concrete cases: the widely accepted connection between DNA and heredity, and the still controversial connection between synaptic plasticity and learning (Morris et al., 2003; Silva, 2003). There are countless experiments connecting DNA to heredity and plasticity to learning; the examples chosen to illustrate these connections do not reflect either historical precedence or importance. Instead, I will simply use a few straightforward and related examples of both cases, with a focus on the least established connection of the two (that between plasticity and learning). Importantly, the *Convergent Four* should be helpful to all disciplines of science, irrespective of phenomenological level (the examples used below span a wide range of complexity levels, from molecular to organismal) or discipline. Thus, this hypothesis assumes that there are no *a priori* restrictions to which two phenomena can be connected other than the ethical and physical limitations imposed by the work required to test the links between them.

4.2. Negative alteration

The Convergent Four hypothesis assumes that all four criteria are equally important for making connections between phenomena in science, and that it is the synergism between lines of evidence in these categories that confers the strength of evidence supporting any one proposed connection.

Negative alteration refers to experiments where the probability of one natural phenomenon is decreased and the effect on another is tested. For example, the alpha calmodulin kinase II gene (aCaMKII) encodes a protein that is present in synapses (points of communication between neuronal cells in the brain) and modulates changes in the efficacy with which certain brain cells communicate (synaptic plasticity). Genetically engineered mutations that delete the aCaMKII gene in mice (negative alteration) lead to a decrease in the probability of synaptic plasticity; inheritance of these mutations results in murine pedigrees that do not have either functional copies of the gene or normal synaptic plasticity. Similarly, since synaptic plasticity is known to have a role in learning, manipulations that decrease synaptic plasticity, such as the mutation of the aCaMKII gene, result in decreases in the probability of learning (Silva, 2003). Together these findings are consistent with the established connections between DNA and heredity (deleting specific genes disrupts specific inherited traits) and between synaptic plasticity and learning (decreasing the probability of synaptic plasticity decreases the probability of learning) (Morris et al., 2003; Silva, 2003). This cross-referencing between connections is critical in science and will be discussed again later.

4.3. Positive alteration

Positive alteration describes experiments where the probability of one natural phenomenon is increased and the influence on another is tested. For example, increasing the number of NMDA receptor 2B genes in the brain results in more NMDA receptor 2B protein and *increases* in the probability of synaptic plasticity (Silva, 2003). This gene encodes a subunit of the NMDA receptor that, among other things, contributes to the activation of CaMKII (Schulman, 2004). Mice that inherit the extra copies of the NMDAR 2B gene express more plasticity in certain brain synapses, and acquire information faster of a type that is known to depend on those synapses (Tang et al., 1999). Thus, these results support both the connection between DNA and heredity, and that between plasticity and learning.

4.4. Non-intervention experiments

Non-intervention refers to experiments designed to determine whether one natural phenomenon follows another. Unlike *positive and negative alteration*, *non-intervention* experiments do not *intentionally* alter the phenomena studied, and as far as there is any experimental interference on the phenomena observed, this is an unwanted and often unintended consequence of the experimental design. Experiments in multiple species and brain systems have documented the o*bservation* that learning is accompanied by changes in synaptic plasticity in the very brain regions required for that particular form of learning (e.g., Moser et al., 1993). Similarly, there are countless *non-intervention* experiments that document the connection between genetic traits and genes. For example, inheriting intact aCaMKII and NMDA receptor genes is critical for inheriting normal synaptic plasticity and learning (Silva, 2003).

It is important to note that non-intervention experiments are also critical to initially define each of the phenomena that are to be connected, but these non-intervention experiments are fundamentally different from those designed to test the connection between phenomena of interest. The first attests to the existence and properties of the phenomena, while according to the Convergent Four hypothesis, the second is one of four strategies to test connections between phenomena of interest. Before connecting genes with heredity, each of these two phenomena (genes and heredity) needed to be observed and defined.

4.5. Integration

Integration refers to efforts in science that do not involve either manipulation or observation of natural phenomena, but are instead focused on proposing, ordering and cross-referencing connections between phenomena. An important component of this activity is the generation of ideas that attempt to explain how one phenomenon is connected to another. For example, Watson and Crick proposed a model for the connection between the biochemical properties of DNA and heredity (Watson and Crick, 1953a, 1953b) that had an enormous impact on the scientific process that connected these two phenomena. Often, one of the two phenomena being connected has never even been observed. For example, Hebb proposed that learning is due to a hypothetical mechanism that modulates the communication between cells in the brain (Hebb, 1949) Mechanisms that modulate cell–cell communication (synaptic plasticity) in the brain were not discovered until more than 20 years later (Bliss and Lømo, 1973). Hebbian-inspired ideas of how synaptic plasticity can be used to store information are an essential part of the body of work that connects synaptic plasticity with learning and memory (Silva, 2003) because they suggested ways of how one leads to the other. Similarly, Watson and Crick's ideas also had the same effect: they showed how the combinatorial possibilities of the bases (A,G,C,T) along a strand of DNA could be used to encode many different genes, and how the complementarity of the two strands of a DNA molecule could account for the propagation of genes between cell divisions and therefore, the transmission of traits between generations. It is important to note that *Integration* efforts do not always precede efforts in the other three strategies to test connections in science. Very often *integration* work accompanies developments propelled by results with the other three approaches.

Another key goal of *Integration* efforts in science is to connect whatever else is known about A and B with a proposed or established connection between A and B. For example, there have been extensive modeling efforts to integrate what is known about plasticity and learning with other knowledge about either of these two phenomena. It is well known that learning depends on arousal, emotion and motivation. Therefore, *Integration* efforts have attempted to connect the biology of these phenomena with the known biology of learning (McGaugh, 2000). This integrative effort leads to additional hypothetical connections that can then be tested with the approaches described above. This is an iterative process that results in the formation of clusters of interconnected connections, where the strength of one connection contributes to the strength of other related connections in a cluster. For example, finding that the *positive alteration* of a biological component of arousal (i.e. beta-adrenergic receptor function) strengthens both synaptic plasticity and learning, and that the negative alteration of the same component weakens both of these two connected phenomena (synaptic plasticity and learning), contributed to the strength of connections in the motivation/learning cluster (i.e., between plasticity and learning, between beta-adrenergic receptors and arousal and between arousal and learning).

Large networks of clusters of connections form fields in science (for example the field of learning and memory), and interconnected groupings of these networks of clusters form disciplines such as Neuroscience. One can imagine all scientific knowledge as residing in a multidimensional space populated by groupings of these networks of connection-clusters, where temporary disciplinary boundaries outline areas where connections between groupings of these networks are scarce. As science advances, this geography of connections changes and so does the taxonomy of the resulting fields and disciplines. The goal of science is to fill this phenomena space with networks of clusters that are optimally interconnected into a seamless and continuous information space. In the distant future, disciplinary boundaries (areas of low density of connections), if any, will be only due to fundamental ethical and physical limitations on experimentation. For example, I hope it will always be unethical to carry out *positive* and *negative alteration* experiments to test links between the properties of sub-atomic particles and the sociology of human populations. Thus, the gaps between particle physics and sociology will be filled by the collective strength of the myriad of interconnected clusters that reside between the information space that separates these two disciplines and by extensive *Integration* efforts that will require considerable amounts of computing power.

4.6. Koch's postulates

The literature includes partial versions and specific instantiations of the *Convergent Four* hypothesis, the most notable of which is commonly referred to as the *Koch (Henle) Postulates* (Koch, 1884; Henle, 1938). These postulates refer to criteria that must be fulfilled before a given pathogen (bacteria, virus) could be proven to *cause* a specific infectious disease, such as cholera, tuberculosis, etc. The four requirements include (1) consistent isolation of the agent from disease cases, but not from normal subjects (this second part was later abandoned by Koch because he found asymptomatic carriers of cholera), (2) propagation of the agent in culture, (3) reproduction of the disease by the isolated purified agent, (4) re-isolation of the agent from cases in "(3)". It is easy to see the commonalities between *Koch's Postulates* and two of the experimental categories described in the *Convergent Four* hypothesis. Indeed, the first, second and fourth of Koch's Postulates, listed above, are specific instantiations of the *Non-intervention* criteria in the *Convergent Four* hypothesis: the agent must be observed in association with the disease, and must be seen to reproduce in appropriate media. The third postulate, which says that the isolated agent must bring about the disease, is a specific implementation of the *Positive Alteration* strategy described in the *Convergent Four* hypothesis (if A and B are connected, then increasing the probability of A should increase the probability of B).

The *Convergent Four* hypothesis predicts that Koch's postulates are not sufficient to compellingly connect a given agent with a specific disease, a prediction underscored by several revisions of Koch's Postulates specifically proposed to address its experimental limitations (e.g., Hill, 1965; Evans, 1976). The *Convergent Four* hypothesis implies that in addition to Koch's Postulates (1) it would be critical to show that deleting the infectious agent from the subject (i.e. use of antibiotics that eliminate the pathogen) cures the disease (*Negative Alteration*), and that (2) it is essential to have a credible explanation for how the pathogen causes the disease. Indeed, a recent revision of *Koch's Postulates* proposes a number of changes that include specific instantiations of both the *Negative Alteration* and *Integration* requirements proposed by the *Convergent Four* hypothesis (Fredericks and Relman, 1996).

Although each causal connection in science requires particular experiments and it is subject to specific experimental and ethical limitations, the *Convergent Four* hypothesis provides an outline for the kinds of experiments and evidence required for a compelling and efficient test of the hypothesis. It is conceivable that it would be possible to make a compelling case for a connection between two natural phenomena even in the absence of one of the experimental categories listed in this hypothesis (i.e., due to insurmountable ethical or technical limitations), perhaps by both collecting extensive amounts of evidence in the other three categories of experiments outlined in the *Convergent Four* hypothesis, and by strengthening the experimental support for other connections related to that connection; causal connections in science are never isolated but within a network of other related connections, suggesting that strengthening related connections can strengthen any one given connection. Perhaps, network theory analysis could be used for a formal analysis of this process (Newman, 2003).

4.7. Tool development

As stated above, the *Convergent Four* hypothesis does not pertain to the other two key goals of experimentation in science, including to define individual phenomena of interest, and the development of new strategies and tools to both define new phenomena and test connections between them. Tool development is a critical component of every one of the four strategies described in the *Convergent Four* hypothesis. For example, neural network models have been an important tool in exploring how synaptic plasticity affects learning, just as X-ray crystallography (still a critical biochemical tool today) was critical to build the structural models that played such a big role in connecting DNA with heredity (Watson and Crick, 1953a, 1953b). Similarly, pharmacological (i.e. NMDA receptor blockers) and genetic tools (mice with CaMKII mutations) were central to efforts to connect synaptic plasticity to learning (Silva, 2003), just as cloning tools were critical to cement the connection between DNA and heredity. Tool development is often dissociated from the very activity of connecting natural phenomena and is an important end in of itself in science. There are numerous examples of how the process of exploring hypotheses in science led to ideas that resulted in very influential tools with wide ramifications in science. For example, Einstein first described the theory of stimulated emission in 1917, which would eventually lead to the development of today's lasers through a series of breakthroughs from people such as Townes, Basov and Prokhorov (development of the mazer), Schawlow and Townes (optical mazer or laser) Alferov and Kroemer (semiconductor heterostructures) Chu, Cohen-Tannoudji and Phillips (laser cooling), etc. Modern lasers are a key tool in a wide variety of scientific areas unrelated to those that originated them. For example, lasers are being used to study synaptic plasticity (Trachtenberg et al., 2002).

4.8. Evaluating the Convergent Four hypothesis

Many of the biologists, especially molecular biologists, with whom I have discussed the *Convergent Four* hypothesis recognize it as part of the implicit *modus operandi* of their discipline. Nevertheless, the underlying assumptions of this hypothesis, as well as its many implications will have to be rigorously tested, just as any other organizing hypothesis in science. S2 is first and foremost a science, and its hypothesis should be tested with the same general strategies as all other hypotheses in science. The process of tool development will be critical for S2, just as it has been critical for every discipline in science. Fortunately, as mentioned above, there are already tools that could be used for testing S2 ideas such as *The Convergent Four*. For example, databases such as Medline can be used to follow the progression and impact of specific experiments on the reliability of ideas in science, and modeling tools such as the Robot Scientist and Discovery Informatics could be used to test the usefulness of these ideas in models of the discovery process.

5. The Convergent Four hypothesis and the study of cognition

Recently, studies of cognition have involved seven main disciplines, including Molecular Neurobiology (i.e., the characterization of key synaptic receptors and downstream molecular-signaling pathways), Neurophysiology (i.e., the study of mechanisms of synaptic plasticity), Systems Neuroscience (i.e., analysis of circuit representations such as place fields), Behavioral Neuroscience (i.e., lesion studies of brain function), Cognitive Neuroscience (i.e., imaging studies of brain function) Neuropsychology (i.e., human subject studies of cognition) Computational Neuroscience (models about cells systems) and Cognitive Science (i.e., models of brain function). All of these disciplines have their own experimental traditions and *modus operandi* that have resulted in biases towards one of the experimental approaches outlined in the *Convergent Four* hypothesis. For example, the majority of studies carried out in Cognitive Neuroscience involve imaging studies of brain activation using techniques such as functional magnetic resonance imaging (fMRI) (Cabeza and Nyberg, 2000). This means that many of the conclusions reached in that field are heavily dependent on *Non-intervention* experiments (activation of brain region X accompanies behavior Y; therefore region X mediates behavior Y). Another neuroscience discipline that has depended heavily on Non-intervention experiments is Neurophysiology, where many of the studies have characterized synaptic neuronal and glial responses under different experimental conditions (see for example, Bliss and Collingridge, 1993). Neurophysiology has also depended heavily on *Negative Alteration* experiments, which use pharmacology and more recently molecular biology tools such as viral vectors and transgenic mice. In these experiments, the function of a specific molecular component is targeted, and the consequences measured with electric or optic physiology tools; the results are then used to evaluate the connection between a given molecular process and a specific aspect of brain physiology (i.e., the role of NMDA receptors in long-term potentiation). Similarly, Behavioral Neuroscience is also very much dependent on *Negative Alteration* experiments, where the function of a brain region is inferred from the behavior of subjects lacking that brain region. In contrast, Computational Neuroscience is focused on *Integration* experiments, where models are used to explore the possible functional properties of molecular, physiological or neuroanatomical mechanisms.

Although each of these disciplines uses a number of complimentary methods, they each have technical limitations and are dependent on experimental traditions that bias studies towards one of the four approaches in the *Converging Four* hypothesis. However, according to this hypothesis, any finding that does not have convergent evidence from all four approaches (*positive*, and *negative alteration*, *non-intervention* and *integration*) would be suspect and should require further work specifically in the other unexplored experimental categories. For example, hypotheses about the function of specific brain regions that are supported by *negative alteration* studies (i.e., lesions) should be explored with *non-intervention*, *positive alteration* and *integration* approaches. Indeed, hypotheses about the function of specific brain region supported by both *non-intervention* (imaging) and *negative alteration* (lesions) experiments usually carry more weight than those supported by only one of these approaches (see for example, Dade et al., 2002).

6. S2 and the future of the cognitive sciences

The drive for integration in neuroscience has motivated researchers to venture into each other's territories and disciplines. However important, this process has raised many fundamental S2 questions. For example, is it possible to connect molecular mechanisms with behavior without compelling insights about the steps in between? If this can be done, what is the evidence supporting this view? If not, what are the arguments against this integration and how can we test these arguments? Is there an *a priori* limitation on which phenomena can be connected? For example, most neuroscientists would agree that it is unlikely that the physical properties of quarks will ever be meaningfully connected to the behavior of societies; what are the formal reasons for this view? Can we say that the properties of NMDA receptors and CaMKII are as tightly connected to spatial learning as the properties of DNA are connected to heredity? If not, why not? How do we know when there is enough evidence to substantiate a connection between brain phenomena? For example, is there sufficient evidence to unequivocally demonstrate that synaptic plasticity is a core mechanism of learning, just as DNA replication is a core mechanism of heredity? If not, what else needs to be added to current evidence and why? What constitutes "sufficient" evidence for any idea in the Biology of Cognition? Systematic, evidence-based attempts to answer these questions, so central to research on cognitive problems, will not only affect our perception of the state and reliability of current knowledge, they will also improve the efficiency with which we explore this immense, but tremendously exciting problem space.

Acknowledgements

This manuscript would not have been possible without the many illuminating discussions with members of my laboratory over the last 15 years. Many thanks go to, Tawnie Silva, Elenna Silva, Alex Silva, Alan Smith, Rusiko Bourtchouladze, Rui Costa, Mehta Mayank, and Matt Glasser for ideas and support that allowed me to talk about S2 and the *Convergent Four* hypothesis without blushing. A special thanks goes to John Bickle, whose knowledge, enthusiasm and imagination were so essential for the articulation and development of the key ideas described in this manuscript. I will be forever grateful to Anthony Landreth for joining our group with the principal goal of helping us develop the ideas outlined here. I am certain that S2 will one-day rest on the shoulders of young thinkers/scientists with his rigor, courage, determination and creativity.

References

Agresti, W.W., 2003. Discovery informatics. Commun. ACM 46, 25–28.

Bliss, T.V., Collingridge, G.L., 1993. A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361, 31–39.

Bliss, T.V., Lømo, T., 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J. Physiol. 232, 331–356.

Cabeza, R., Nyberg, L., 2000. Imaging cognition II: an empirical review of 275 PET and fMRI studies. J. Cognitive Neurosci. 12, 1–47.

Cole, J., 2000. A short history of the use of citations as a measure of the impact of scientific and scholarly work. In: Cronin, B., Atkins, H. (Eds.), The Web of Knowledge. Information Today Inc., Medford, New Jersey, pp. 281–300.

Cole, J., Cole, S., 1972. The Ortega hypothesis. Science 178, 368–374.

Dade, L.A., Zatorre, R.J., Jones-Gotman, M., 2002. Olfactory learning: convergent findings from lesion and brain imaging studies in humans. Brain 125, 86–101.

Ebbinghaus, H., 1885. Über das Gedchtnis. Untersuchungen zur experimentellen Psychologie. Duncker & Humblot, Leipzig.

Elgersma, Y., Silva, A.J., 1999. Molecular mechanisms of synaptic plasticity and memory. Curr. Opin. Neurobiol. 9, 209–213.

Evans, A.S., 1976. Causation and disease: the Henle–Koch postulates revisited. Yale J. Biol. Med. 49, 175–195.

Fredericks, D.N., Relman, D.A., 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. Clin. Microbiol. Rev. 9, 18–33.

Hamilton, D.P., 1990. Publishing by – and for? – the numbers. Science 250, 1331–1332.

Hamilton, D.P., 1991. Research papers: who's uncited now? Science 251, 25.

Hebb, D.O., 1949. The Organization of Behavior. John Wiley and Sons, New York.

Henle, J., 1938. On Miasmata and Contagie. Johns Hopkins Press, Baltimore.

Hill, A.B., 1965. The environment and disease: association or causation? Proc. R. Soc. Med. 58, 295–300.

Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. Proc. Natl. Acad. Sci. USA 102, 16569–16572.

Kanigel, R., 1993. Apprentice to Genius: The Making of a Scientific Dynasty. Macmillan, New York.

King, R.D., Whelan, K.E., Jones, F.M., Reiser, P.G., Bryant, C.H., Muggleton, S.H., Kell, D.B., Oliver, S.G., 2004. Functional genomic hypothesis generation and experimentation by a robot scientist. Nature 427, 247–252.

Koch, R., 1884. Die Aetiologie der Tuberkulose. Mitt Kaiser Gesundh 2, 1–88.

Lavoisier, A., 1789. Traité élémentaire de chimie, présenté dans un ordre nouveau et d'après les découvertes modernes. Chez Cuchet, Paris.

Martin, S.J., Grimwood, P.D., Morris, R.G., 2000. Synaptic plasticity and memory: an evaluation of the hypothesis. Annu. Rev. Neurosci. 23, 649–711.

McGaugh, J.L., 2000. Memory – a century of consolidation. Science 287, 248–251.

Merton, R.K., 1968. The Matthew effect in science. Science 159, 56–63.

Morris, R.G., Moser, E.I., Riedel, G., Martin, S.J., Sandin, J., Day, M., O'Carroll, C., 2003. Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory. Philos. Trans. R. Soc. Lond. B, Biol. Sci. 358, 773–786.

Moser, E., Moser, M.-B., Andersen, P., 1993. Synaptic potentiation in the rat dentate gyrus during exploratory learning. Neuroreport 5, 317–320.

Newman, M.E.J., 2003. The structure and function of complex networks. SIAM Rev. 45, 167–256.

Pavlov, I.P., 1927. Conditioned Reflexes. Oxford University Press, London.

Polya, G., 1945. How to Solve It: A New Aspect of Mathematical Method. Princeton University Press, Princeton.

Rose, S.P., 1981. What should a biochemistry of learning and memory be about? Neuroscience 6, 811–821.

Schulman, H., 2004. Activity-dependent regulation of calcium/calmodulin-dependent protein kinase II localization. J. Neurosci. 24, 8399–8403.

Schunn, C.D., Dunbar, K., 1996. Priming, analogy, and awareness in complex reasoning. Mem. Cognition 24, 271–284.

Shackelford, B., 2006. U.S. R&D continues to rebound in 2004. In: InfoBrief Science Resources Statistic. National Science Foundation, Directorate for Social, Behavioral and Economic Sciences, Washington, DC, pp. 1–6.

Silva, A.J., 2003. Molecular and cellular cognitive studies of the role of synaptic plasticity in memory. J. Neurobiol. 54, 224–237.

Smith, A., 1776. An Inquiry into the Nature and Causes of the Wealth of Nations. W. Strahan and T. Cadell in the Strand, 1776, London.

Tang, Y.P., Shimizu, E., Dube, G.R., Rampon, C., Kerchner, G.A., Zhuo, M., Liu, G., Tsien, J.Z., 1999. Genetic enhancement of learning and memory in mice. Nature 401, 63–69.

Tansey, M., Stembridge, B., 2005. The challenge of sustaining the research and innovation process. World Patent Inform. 27, 212–226.

Trachtenberg, J.T., Chen, B.E., Knott, G.W., Feng, G., Sanes, J.R., Welker, E., Svoboda, K., 2002. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature 420, 788–794.

Watson, J.D., Crick, F.H., 1953a. Genetical implications of the structure of deoxyribonucleic acid. Nature 171, 964–967.

Watson, J.D., Crick, F.H., 1953b. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171, 737–738.

Whelan, K.E., King, R.D., 2004. Intelligent software for laboratory automation. Trends Biotechnol. 22, 440–445.

Wren, J.D., Bekeredjian, R., Stewart, J.A., Shohet, R.V., Garner, H.R., 2004. Knowledge discovery by automated identification and ranking of implicit relationships. Bioinformatics 20, 389–398.

Zuckerman, H., 1967. Nobel laureates in science: patterns of productivity, collaboration and authorship. Am. Sociol. Rev. 32, 391–403.