Printed in: *New Ideas* in *Psychol,* Vol. 8. No.2, pp, 231-246, 1990 0732-118X/9O Printed in Great Britain @ 1990 Pergamon Press

SECRETS OF GENIUS

Review of *Imagery in Scientific Thought* by Arthur I. Miller. Cambridge, MA: MIT Press, 1986.

DAVID HESTENES

Department of Physics, Arizona State University, Tempe, AZ 85287-1504, U.S.A.

Scientists and nonscientists alike are fascinated by the creative processes underlying the great scientific discoveries. We are eager to know the secrets of genius. Did Einstein possess creative powers that set him above the ordinary physicist? Or was he privy to some special heuristics that guided him to his discoveries? We are indebted to historians of science like Miller for helping us answer such questions. Recognizing the difficulty of the task, Miller calls for collaboration between historians and cognitive scientists to study creative processes in science. He tries to get the process started in the present book with a historical, epistemological, and cognitive analysis. His central thesis is that "mental imagery is a key ingredient in creative scientific thinking." We follow him by focusing attention on the role of imagery in the creation of the special theory of relativity and quantum mechanics, two major triumphs of 20th century physics. But to evaluate the role of imagery we need to know what else was involved in the creation of these great theories.

THE CREATION OF SPECIAL RELATIVITY

Miller previously wrote an entire book on the creation of special relativity theory. The present volume summarizes and extends that excellent work. Let me try to put the whole story in a nutshell. Though special relativity is often said to be a revolutionary theory, Einstein himself regarded it as a straightforward extension of classical physics. It does no more than correct a deficiency in the classical concept of time. The deficiency came from taking the concept of simultaneity for granted. To analyze the concept, Einstein devised a *gedanken* (or thought) experiment for determining when spatially separated events are simultaneous. This revealed an arbitrariness in the concept of simultaneity that had hitherto gone unnoticed. It enabled him to prove that, if signals used in measurements propagate with finite velocities, then judgments of simultaneity depend on the chosen reference frame. More specifically; if events A and B are simultaneous in one reference frame, then there exist other reference frames in which A occurs before B or B occurs before A. This *relativity of distant simultaneity* is the core of Einstein's new insight into the concept of time. It is this that justifies the name "relativity" for the theory.

To understand the origins of special relativity theory is to recognize that creation of the theory was an inevitable consequence of the scientific enterprise. Einstein himself declared that someone else would surely have done it if he had not. Need for the theory arose from an incompatibility between the two great theories of classical physics: Newtonian mechanics and the newly completed electromagnetic theory of Faraday and Maxwell. This incompatibility manifested itself in a variety of paradoxical experimental results, so it could hardly be overlooked or ignored. Indeed, it engaged the attention of many outstanding researchers, led by H. A. Lorentz and Henri Poincaré. How was it that Einstein alone was able to identify the locus of the problem in the concept of simultaneity?

Miller identifies a number of ideas and observations that are likely to have helped guide Einstein's thinking. We note here two of the major ones which, by themselves, are nearly sufficient to determine the two basic postulates of relativity theory.

Einstein did not need an elaborate analysis of experimental data to identify the conflict between Newtonian mechanics and electromagnetic theory. Both theories are involved in explaining the phenomenon of electromagnetic induction, which underlies the operation of electric motors and generators. The essence of the phenomenon is that a magnet moving relative to a wire loop induces an electric current in the wire. Einstein observed that the induced current predicted by the theory depended on whether the wire or the magnet was kept at rest, whereas the physical phenomenon appeared to depend only on the relative motion of the magnet and wire. Thus, the theory exhibited an asymmetry which was not inherent in the phenomena. Einstein removed this asymmetry by invoking the *principle of relativity,* which requires that *the laws of physics for an observer at rest must be the same as for an observer moving with uniform velocity.* This principle had been stated for mechanics by Newton, though not as a basic axiom. Einstein generalized it to apply to electromagnetic theory as well. Paradoxically, this required a modification of mechanics rather than electromagnetics. The precise form of the modification was determined by the second postulate of Einstein's theory.

Einstein was evidently led to his second postulate by pondering a *gedanken* experiment in which the speed of an observer is increased until it equals that of a travelling light wave under observation. If this were possible, then the light wave would appear as a standing wave. But no such standing wave has ever been observed in optical experiments; therefore, it must be impossible. In accordance with the relativity principle, Einstein excluded this possibility and assured that the laws of optics will be the same for all observers by postulating *that the speed of light is the same for all observers and independent of the velocities of bodies emitting light.*

After Einstein had settled on his two basic postulates he was able to derive the necessary modification of Newtonian mechanics mathematically. That involved deriving an explicit mathematical expression for the transformation relating the space and time coordinates of one observer to those of another. The result is now called a *Lorentz transformation,* because Lorentz had, in fact, derived it before Einstein. However, the physical meaning of the transformation in Einstein's theory is profoundly different from its meaning in Lorentz's theory. This point is not sufficiently emphasized in Miller's account. The physical interpretation of the mathematics is a crucial ingredient of a physical theory. It takes more than the mathematical form of the Lorentz transformation to arrive at the fundamental underlying concept, namely, the relativity of distant simultaneity. Here it may be surmised that Einstein once again made use of the symmetry considerations so prominent in his thinking. He must have been distressed by the nonsymmetrical form of the Lorentz transformation relating the time variables in different reference frames, for the relativity principle demands that both variables be equally significant. To retain the symmetry in the interpretation of the theory, he may well have explored *gedanken* experiments relating time measurements in the different reference frames. That would have led him to the fundamental insight that different reference frames necessarily entail different definitions of simultaneity. Too bad he is not around to ask!

We have here a rational reconstruction of how Einstein might have identified simultaneity as the crucial issue. The greatest remaining mystery is why Poincaré failed to arrive at the same conclusion and, indeed, to appreciate Einstein's accomplishment in subsequent years. Miller shows us that Poincaré was well aware of all the essential facts and ideas. The only thing missing, it seems, was an appreciation of *gedanken* experiments.

This case illustrates an important difference between mathematical and physical thinking which goes a long way toward explaining why so few mathematicians have made important contributions to physics in the 20th century. Pure mathematicians do not think about the equations of physics in the same way as a physicist does. They are concerned only with the structure of the equations and the formal rules for manipulating them. But physicists regard the equations as representations of real things or processes; they are only partial representations of the physicists' knowledge, so to improve a representation they may alter the equations in ways that violate mathematical rules. Both Einstein and Heisenberg were masters at this. Neither was a mathematical virtuoso. Indeed, in the period when Einstein was developing his general relativity theory, the mathematician Hilbert expressed the opinion that Einstein was mathematically naive. I have heard a similar opinion about Heisenberg expressed by one of his students in later years. Mathematics played an essential role in Einstein's thinking, but, as mathematical physics goes, the mathematics in all his great papers is comparatively simple. His forte was in analyzing the physical meaning of the mathematics. Indeed, such analysis is generally characteristic of the best work in theoretical physics. I have heard the Nobel laureate Richard Feynman, himself a true mathematical virtuoso, express this opinion forcefully, asserting that the value of a paper on theoretical physics is inversely proportional to the density of mathematics in it.

What secrets of Einstein's genius have been revealed by studying the genesis of relativity theory? Consider the following four prominent characteristics of his thinking.

(1) *Skepticism and objectivity*

A successful scientist no less than a successful writer must be his own best critic. The crank is distinguished from the scientist not so much by his outlandish ideas as by a refusal to subject his favorite ideas to a searching criticism. Einstein's admiration for Mach's "incorruptible skepticism" is self-revealing. His own skepticism was expressed in the way he zeroed in on his own intellectual presuppositions as well as those of others. He never let his ego interfere with the objectivity of his scientific evaluations. Thus he was able to recognize and assimilate valuable ideas from other thinkers, like Mach, with whom he had profound disagreements. In ironic contrast, Mach was never able to appreciate Einstein, repudiating relativity to his death.

(2) *Search for general principles*

This is the constructive complement to Einstein's critique of presuppositions, for the generic laws and principles of physics are among the strongest presuppositions. Einstein never engaged in the data-driven search for specific laws by curve-fitting to experimental data. Rather, he looked for simple principles providing broad constraints on the range of possible empirical results. By itself, such a principle predicts nothing specific, but in combination with other principles it may have unlimited implications. In his insistence that such principles are "freely invented," Einstein avoided the extremes of empiricism and rationalism. He emphasized that the basic principles of physics are not summaries of or abstractions from experience, nor can they be found by deductive argument. This is not to say that they cannot be discovered (or should I say invented) by rational means. His own discovery of the principles of relativity theory shows otherwise.

(3) *Symmetry arguments*

Such arguments have been used informally in physics from its beginning, but never before so explicitly and effectively as by Einstein. He was the first to incorporate a symmetry principle into the very foundations of physics. Inspired by the striking success of relativity theory, physicists have developed symmetry arguments into a systematic methodology with a formal structure provided by group theory, the mathematical theory of symmetry. Group theory has become one of the most powerful tools of theoretical physics, in daily use for manifold applications of established theory as well as the development of new theory. Thus, one of the major heuristics in Einstein's great discovery has become a standard tool of ordinary physicists. It is interesting to note that group theory had been formalized only a short time before relativity theory. Since Einstein's theory can be regarded as one of the greatest applications ever of group theory, one wonders if it was at all influenced, perhaps indirectly, by the formal mathematical theory.

(4) Gedanken *experiments*

Such experiments were used to assess the semantic content (i.e., the physical meaning) of theoretical constructs. Though Einstein did this with consummate skill and had the privilege to employ it in some of the greatest discoveries in history, it is not so unique to Einstein as Miller suggests (p. 113). It should be recognized as one of the basic mental activities of any "real physicist." Indeed, it could be regarded as a defining difference between physicist and mere mathematician. Galileo and Newton did it with equal skill. Mach's provocative positivist critique of Newtonian mechanics must have helped Einstein develop his own skill. Surely the design of any physics experiment begins with a conceptual dry run; so, Faraday must be counted as one of the all time experts in *gedanken* experiments. Likewise, the major role of experimental design in the thinking of Enrico Fermi (1950) can be seen in problems he assigned to students in his unique textbook on nuclear physics. So it goes.

The thinking in Einstein's creation of relativity theory can be described as *theory-driven.* As we have seen, it was not directed at explaining any particular experimental results, but it was nonetheless empirically grounded in a broad and indirect way. This made empirical predictions from the theory exceptionally robust. As Miller explains (p. 118), the empirical data available in 1901 contradicted Einstein's theory as well as Lorentz's theory of electrons. Since Lorentz's theory was *data-driven,* he was ready to abandon it immediately in deference to the new data. But the rationale for Einstein's theory was so secure that he confidently dismissed the data as inaccurate. Strong empirical confirmation for relativity theory was not available for decades. Nevertheless, many physicists came to accept it on the basis of its internal logic.

The more we know about the invention of relativity theory, the more we see it as a rational process, no less admirable for having the mysteries of its origin removed. The characteristics of Einstein's thinking which we have identified are similar in kind (if not in degree) to those of any theoretical physicist. They all involve cognitive skills that develop only with concentrated practice. Like every other creative scientist, Einstein was intensely dedicated to his work. And far from the myth of the solitary genius, he learned from the best teachers, honing his skills in thorough study of the works of such master physicists as Mach, Boltzmann, and Lorentz. The sharp critical skills he developed early in life enabled him to move quickly to a study of the most productive issues. From Einstein's example there is still much to be learned about the training of our young scientists.

SPACETIME AND THE RELATIVITY SCANDALS

Einstein's original 1905 article determined the content of special relativity theory completely, and no modifications of the theory have been made since. However, it took some time thereafter to recognize the deeper implications of the theory. The main advance came from a reformulation of relativity theory due chiefly to H. Minkowski. It consists of (a) a new mathematical form for equations of physics now called the *covariant form* along with (b) a new diagrammatic method for representing physical events now called *Minkowski* (or *spacetime)* diagrams. Einstein has asserted that Minkowski's inventions were absolutely essential to the creation of his general theory of relativity. Miller chose not to discuss them in this book, probably because he was not dealing with general relativity. But they have proved to be as important to understanding as to applying special relativity. Indeed, they have become standard conceptual tools in the field. The reasons for this bear heavily on Miller's central thesis, so some discussion is in order.

A covariant formulation of the equations of physics has the great advantage of eliminating from the equations the imprint of a chosen reference frame. It is thus one step closer to an observer-independent description of physical reality. At the same time, it simplifies the form of the equations as well as many calculations.

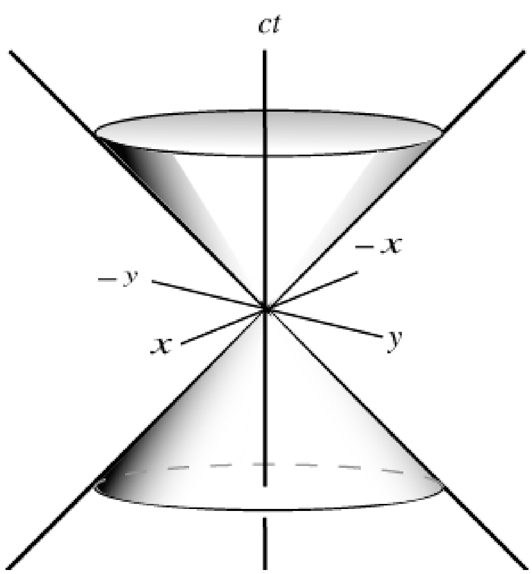

Figure I. A Minkowski diagram of spacetime depicting two dimensions of space and one of time. The spacetime points designate unique events. The lightcone with vertex at the origin consists of all events which can be connected to the origin by a light signal or any other signal travelling in a straight line with speed *c*. The time axis shown is not unique; any other axis within the lightcone will serve as well, each one representing the time coordinate for some (inertial) reference frame.

More important, it reveals a deeper physical unity, including the unification of energy and momentum conservation laws into a single law. Underlying this is a new conception of space and time inherent in relativity theory but not initially recognized by Einstein. Minkowski's new mathematical and diagrammatic representations brought that conception to the fore. The old notions of three-dimensional space and one-dimensional time were replaced by the concept of a single four-dimensional *spacetime* manifold in which there is no unique time or space dimension. With the help of Minkowski diagrams (Figure 1) physicists have become adept at visualizing physical processes in the four-dimensional spacetime arena and have thus developed a new kind of "physical intuition." Though no one has a four-dimensional sensorium, mental imagery in four dimensions can be developed in much the same way as in three. As Boltzmann explains it, "the idea of a three-dimensional figure has no content other than the ideas of a series of visual images which can be obtained from it, including those which can be produced by cross-sectional cuts…it is held together or united only by a clear idea of the laws in accordance with which its perspective images follow one another" (Miller, 1986, p. 49)

The spacetime conception provides new insight into the significance of the speed of light constant *'c'* so prominent in Einstein's second postulate. Note that in Figure 1 the constant plays the role of a conversion factor changing the time variable *'t'* into a variable *'ct'* with the same unit as the space coordinates *'x'* and *'y.'* A common unit like this is just what is needed for the conception of spacetime as homogeneous with no preferred time or space directions. Thus, the speed of light (and with it the spacetime lightcone) is seen to represent a generic property of spacetime rather than a mere special property of light. Today we know that other entities, such as neutrinos, travel with the same speed, c, which is to be expected if the lightcone is indeed an intrinsic property of spacetime.

The crucial role of imagery in understanding relativity is brought out by examining common misunderstandings of the theory. For decades the entire physics community was guilty of promulgating false and misleading explanations of relativistic phenomena to students and the public. No wonder some skepticism persisted even among physicists! The errors were scandalously elementary and egregious. Two scandals, that were not exposed until after Einstein's death, are documented in American Association of Physics Teachers (1963). The crux of each scandal can be explained in a few words.

Relativity theory implies that the measured length L of a rod moving with velocity v along the direction of its axis will be less than the length L_0 measured when the rod is at rest, as given by

$$L = L_0 \sqrt{1 - v^2/c^2}$$

This is called the "Lorentz contraction," because Lorentz wrote down the formula before Einstein. The very name suggests the confusion of more than a few physicists who failed to realize that the respective meanings given to the formula by Lorentz and Einstein were incompatible. For Lorentz the formula described a physical contraction of the rod in the direction of its motion, whereas for Einstein it merely related measurements performed in different ways. However, the scandal was due to physicists who had that much straight. They told everybody that a physical object would actually *appear* to shrink in the direction of its motion. In doing so they confounded the elementary distinction between "a measurement of length" and "visual appearance." This blunder persisted until it was corrected in 1959 by a graduate student James Terrell (ibid.). He pointed out that the light in an instantaneous visual image must have left different parts of the object at different times, because they are different distances away. He showed that when this was taken into account, to a first approximation it had the effect of cancelling the Lorentz contraction so the moving object did not appear distorted.

The second scandal appears to have been a misleading explanation of the twin paradox (also called the clock paradox). It was initiated by Einstein himself (1948) and spread abroad in an influential textbook by Moeller (1972). Although the explanations by Einstein and Moeller may not be literally wrong, they contain irrelevant considerations that make them unnecessarily complex and easy to misinterpret. They encouraged the misconception that "general relativity is needed to deal with accelerated systems." It is of interest to note that all this is related to a complex of misleading statements by Einstein concerning (1) the equivalence of gravitational forces with accelerated systems, (2) the origin of centrifugal forces (or Mach’s principle), and (3) the role of covariance in distinguishing special and general relativity. On the last point Einstein was called to task by Kretchmann (1917). Each of these points have been the subject of much discussion and debate in the physics community. The issues involved in the first and third points are now well understood by the experts, thought they continue to cause problems for novices. The second point involves fundamental physical issues that have not yet been resolved.

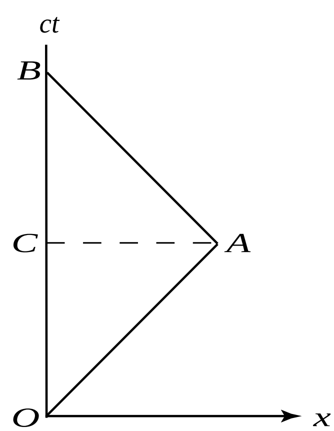

Figure 2. Minkowski diagram for the twin paradox.

Einstein’s statement on these points suggests that he may have harbored some misconceptions of his own about relativity. Even so, he insisted that these ideas played indispensable heuristic roles in the development of general relativity. If he did have misconceptions, he never let them interfere with the objective formulation of his theories. This raises a set of historical and conceptual issues that require the skills of a historian like Miller to be resolved.

Minkowski diagrams played an important role in clearing up the second scandal (American Association of Physics Teachers, 1963) In the twin paradox, one twin in a pair of identical twins stays at home while the other twin travels at high constant speed to a nearby star and then returns immediately. Relativity theory predicts that the return twin will be younger than the stay-at-home twin.

Without discussing the vicissitudes of the paradox, we note that the situation is simply represented by a triangle in Minkowski space (Figure 2). Side OCB of the triangle represents the history (or path in spacetime) of the stay-at-home twin. The other two sides represent the history of the traveling twin, who leaves home at O, turns around at A and arrives back home at B. The length of each side is a measure of the time elapsed in travelling along it. Therefore, computation of the twin age difference after the trip is reduced to “solving the triangle” in Figure 2 for the lengths of the sides. This is no more difficult than solving a triangle in elementary trigonometry. The only difference is that the rules of Euclidean geometry for associating numbers with line segments must be replaced by rules of non-Euclidean geometry. In particular, the Pythagorean theorem for the lower right triangle with vertex at C in Figure 2 is replaced by the non-Euclidean rule, $(OA)^2 = (OC)^2 - (CA)^2$. As a consequence of this change, the bent path OAB is actually shorter than the straight path OCB, and, for spacetime histories in general, the longest path between two events is a straight line.

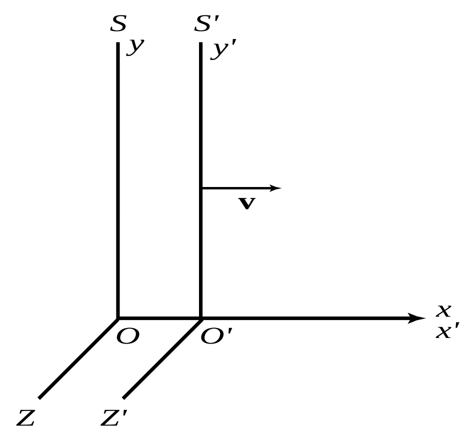

Figure 3. A misleading diagram relating coordinate systems (or frames) moving with relative velocity *v*.

All the elementary problems of relativity involving "time dilations," "Lorentz contractions," and the like, can be reduced to solving triangles in Minkowski space. The simplicity and clarity of this geometrical representation should be compared with more conventional representations of motion in three-dimensional space, where even the most innocent looking diagrams often contribute to confusion by suppressing important distinctions. For example, consider Figure 3 which is supposed to represent a coordinate system S’ moving, with respect to a coordinate system S, with velocity *v* along the *x*-axis. This figure is taken from Moeller (1972), and a similar one appears in most books on relativity. It is dangerously misleading in two respects. First, it invites the viewer to picture S’ as an extended body contracted in its direction of motion. Of course, this is the original conception of Lorentz which becomes a misconception in Einstein's theory. Some books reinforce this misconception by drawing an extended body in the figure or mentioning that the scales on the *x* and *x’* coordinates are different. The second problem with Figure 3 is its ambiguity with respect to time. A coordinate system is commonly regarded as an imaginary extended body with the coordinates tacitly representing points of the body at the same time. With this interpretation Figure 3 is inconsistent with the relativity of distant simultaneity, for if *x* and *x’* are related by a Lorentz transformation, then they cannot represent simultaneous events in both coordinate systems. This example suggests that conventional (Newtonian) representations of motion in three dimensions are more likely to produce confusion than enlightenment in relativistic problems. Historically, there has been a steady trend to replace them by Minkowski diagrams in physics textbooks. The confusion remains mostly at the elementary level.

IMAGERY IN QUANTUM MECHANICS

Miller devotes one chapter to the role of imagery in the development of quantum mechanics. This is a difficult subject, because the history of quantum mechanics is extremely complex, involving many actors. We should not be surprised or disappointed that he is hardly able to do more than get the analysis started. He wisely confines most of his attention to a single individual, Werner Heisenberg. Even so, his chronicle of ideas and events is so compressed (with so many subtleties and caveats omitted) that much of it must be incomprehensible to anyone other than a professional physicist.

Miller notes two major shifts in the imagery of quantum mechanics: from planetary atomic models to energy levels diagrams, and from thence to "Feynman diagrams." These shifts are so prominent in the literature as to be obvious to physicists, but that does not make them any less worthy of a cognitive and epistemological analysis. Let us follow Miller in discussing these shifts, without forgetting that there are other modes of imagery employed in quantum mechanics. Heisenberg was a major actor only in the first shift. He began his graduate research under Max Born, one of the leading practitioners of the "old quantum theory," at a time when the theory was failing badly. Thus, he was in an ideal strategic position to contribute to the field. In the old quantum theory, the states of an electron in an atom were represented as orbits around the atomic nucleus similar to the orbits of planets around the sun. Each orbit has a definite energy, so the various atomic states can be represented as energy levels as in Figure 4.

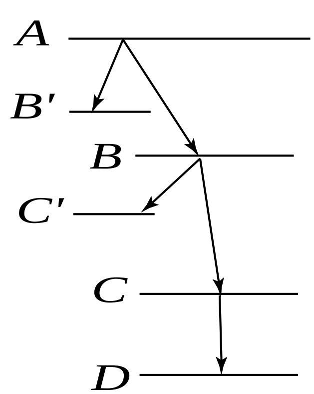

Figure 4. An energy level diagram depicting transitions between energy levels (or states) in an atom. (Taken by Miller from an article by Weisskopf and Wigner, 1930.)

Absorption of light by an atom can then be represented as a transition or jump up to a higher energy level, while emission is represented by a jump down. In the revolutionary change from the old quantum theory to the new, the concept of energy level was retained while the planetary orbits were replaced by a more abstract concept of atomic state. Energy level diagrams evidently played a key role in the thinking of Heisenberg and others during this revolution by providing a representation of atomic processes which is free from the defects of old quantum theory.

Miller gives us a careful analysis of the language used by Heisenberg in his search for a suitable imagery for quantum mechanics. However, this does not provide us with a clean separation between the psychological and epistemological issues involved. Let's see if we can help pry them apart.

The scientists in Miller's account are unanimous in emphasizing the crucial role of visualization in scientific thinking along with a warning that it can be misleading. One place they were misled (along with Miller and the physics community at large) was in their intuition that classical mechanics describes what is perceptually given. They were unaware of the strong cognitive component in their own perception. It was only by training that classical mechanics came to be integrated into that perception. Cognitive research has recently established that the perceptions of people untutored in physics are naturally inconsistent with classical mechanics in almost every detail (Halloun & Hestenes, 1985). Thus, Miller's conclusion (p. 261) that "twentieth-century physicists were forced to liberate their thinking from the world of perceptions" misses the mark. The intellectual struggle to distinguish the objective properties of things from the perceptually given was as crucial to creating classical mechanics as quantum mechanics. Among all the perceptually salient qualities, the founders of classical mechanics had to identify motion as fundamental and provide it with an abstract mathematical representation. The reasoning of Galileo, Newton, and Boyle on the distinction between perception and reality (Burtt, 1980) has not been surpassed in subtlety by Heisenberg and Bohr, though few physicists are familiar with it.

Having recognized the psychological tendency of physicists to confuse classical physics with perception, we can see more clearly the central epistemological issue raised by the creation of quantum mechanics. The conflict between classical and quantum physics had nothing to do with perception. It arose because physicists were unable to reconcile the mathematics of quantum mechanics with the classical conception of reality, so they were forced to construct new "quantum mechanical" conceptions of reality. This involved constructing new images of reality based in large part on diagrams expressing the mathematical structure of the theory. Thus we have a certain incompatibility between the classical and quantum images of reality. This is paradoxical because one cannot simply choose one theory over the other; both theories are essential to physics. The paradox troubled Einstein deeply, and many physicists today do not believe it has found a satisfactory resolution.

Miller's account of the imagery shift from energy level diagrams to Feynman diagrams accomplishes little beyond noting that the shift happened, and it is misleading in some respects. He is certainly right, however, in emphasizing its importance, so it should be worthwhile to expand on his beginnings. Miller's failure to find a direct link from Heisenberg to Feynman is understandable, because there isn't any. Feynman has publicly stated that his diagrammatic methods evolved from his path integral approach to quantum mechanics, which got its original impetus from Dirac's famous textbook on quantum mechanics. His was an idiosyncratic approach to a widely recognized problem in the field of quantum electrodynamics (QED), which treats the interaction of light with matter. The problem was to develop a covariant method for solving the equations of QED. The widespread recognition of the problem is shown by the fact that it was solved independently by Schwinger and Tomanaga, who shared the 1965 Nobel Prize for that with Feynman. It should be understood that this achievement was entirely mathematical in the sense that no alteration of the basic QED equations was required. The basic equations and the standard noncovariant method for solving them were ably expounded by Heitler (1954) in a book known to every physicist in the field. The drawback of the noncovariant method was that once one got beyond the simplest cases it became unwieldy and physicists lost their way in the complexities. The new covariant methods brought great simplifications that made possible the formulation and circulation of subtle experimental effects, such as the Lamb shift, which were verified with unprecedented accuracy.

The covariant methods are still very complicated, so Feynman developed a diagrammatic technique to simplify and systematize their application. Miller mentions only the simplest example of a Feynman diagram shown in Figure 5. He fails to mention that the chief function of diagrams is mnemonic. Each diagram represents a particular mathematical expression, and Feynman developed a system of rules for translating diagrams into mathematics and vice versa. The diagrams are much easier to generate, manipulate and analyze than the mathematics. In the end, however, the computations must be carried out mathematically. Here also Feynman introduced a number of significant 1mprovements in technique.

Besides their mnemonic value, the Feynman diagrams have a suggestive physical interpretation. Thus, Figure 5 suggests that the electromagnetic force between electrons is mediated by photon exchange, though the mathematics it represents does not literally describe the emission, propagation and absorption of a real photon. A caveat is sometimes introduced by saying that the photon exchanged is virtual rather than real. But the caveat is omitted all too often. There is a strong tendency for physicists to reify the mathematics. On Miller's account this is right in line with Heisenberg's view that we should let the mathematics tell us what is real. One trouble with that view in this case is that the Feynman diagrams generated to solve many problems are not unique, but depend on the particular approximation scheme employed. Thus, the physical interpretation of Feynman diagrams raises serious ontological issues that are yet to be resolved. In the meantime, the "particle exchange" interpretation provides a colorful metaphor, at least, that helps physicists "visualize" and talk about their mathematics.

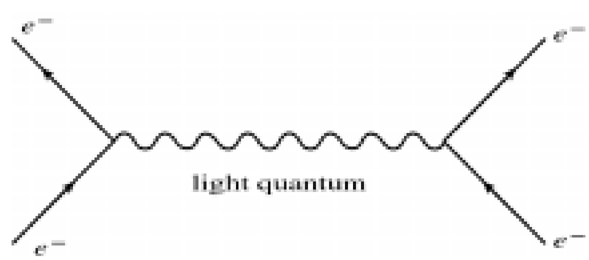

Figure 5. A Feynman diagram depicting the interaction of two electrons by photon exchange.

Miller contrasts the creative thinking of Schwinger and Tomanaga with Feynman's by calling it *non-imaginal.* That won't wash. First, their mathematical creations are not so very different, and, since Dyson showed that they are all equivalent, Feynman's diagrammatic techniques have been fully integrated with the more rigorous mathematics of Schwinger. Second, it may be doubted whether creative thinking in physics is possible without a strong imaginal component. Be that as it may, there is plenty of evidence for imagery in the work of Schwinger and Tomanaga. We have already noted that the imagery of Minkowski diagrams goes hand in hand with covariant equations. Schwinger and Tomanaga were thoroughly schooled in this, and references to spacetime structures in their papers show that it was prominent in their thinking. Though this is not the place to spell out details, it should be noted also that spacetime imagery is essential to the full interpretation of Feynman diagrams. In other words, a Feynman diagram is a kind of generalized Minkowski diagram.

Anyone involved in the lectures, seminars and informal give-and-take of creative physicists cannot fail to notice the vivid imagery in their thinking. Most of this imagery is suppressed in their publications, partly by conventions concerning the style of scientific reporting, partly because it is not essential to establishing the scientific results, and partly because it may be too much trouble to construct suitable diagrams to express it. This puts severe limitations on Miller's historical approach and tells us that the creative physicist needs to be studied in vivo, while he is alive and kicking. That is where the cognitive scientist comes in.

HELP FROM COGNITIVE SCIENCE

Miller devotes the last third of his book to bringing psychology to bear on his theme. This part can be read independently, because it repeats the relevant historical material from earlier chapters. Its deficiencies show how much he needs the help of cognitive scientists.

Miller puts his skills as a historian to good use in setting the record straight on the famous Gestalt reconstruction of the creation of relativity theory by Einstein's good friend Max Wertheimer. He absolves Wertheimer of responsibility for the mistakes of historians and philosophers in misconstruing his reconstruction as a historical account. Unfortunately, Wertheimer's analysis does not give us any new insights into Einstein's thinking. Rather, the detailed historical account of Einstein's creation helps us recognize weaknesses in Wertheimer's reconstruction.

In another chapter, Miller makes a misdirected attempt to apply Piaget's theory of cognitive development to analyze the genesis of relativity and quantum mechanics. In the preface he admits that he was warned that Piaget's three developmental stages (sensori-motor, concrete operational, formal operational) apply only to children, but that did not deter him from applying them in a loosely metaphorical way to the thinking of mature physicists. It is time he learned that there is a large literature on the applicability of Piaget's theory to learning and teaching physics (Hestenes, 1979, 1987). It is well established that the highest cognitive level in Piaget's theory, the formal operational stage, is essential to understanding physics, though, contrary to Piaget's original evidence, many adults never reach it. Piaget may have envisaged extensions of his theory to account for the higher creative activities of science, but according to Piaget's analysis the creative powers already present at the formal level are so rich that perhaps no qualitatively higher level exists.

Turning to cognitive science for more recent research on the role of imagery in scientific thought, Miller gets distracted by an AI dispute about the possibility of propositional representations for mental images and emerges without helpful insights. The deepest insights are to be found in the work of Herbert Simon, who aims to take the mystery out of creativity in science. As a working hypothesis for cognitive research, he makes the strong claim that the creative process in science is nothing more than problem solving, not differing from ordinary problem solving in science, except occasionally in the use of more powerful heuristics. To confirm this hypothesis, Simon has accepted the challenge to exhibit significant heuristics explicitly. In collaboration with others, he has formulated a number of heuristics with sufficient precision to produce computer programs capable of nontrivial problem solving. There is no obvious limit to how far such research can be carried. With Miller's help we have already identified powerful heuristics used by Einstein. Some of these heuristics—for example, symmetry arguments—have become standard conceptual tools in physics.

Miller underestimates the possibilities for analyzing imagery by Simon's approach (p. 261). The analysis of perception in chess by Chase and Simon (1981) provides sharp insights into the workings of perception in high level cognitive skills which need not be reviewed here. Their analysis suggests that Poincaré's ability to grasp a mathematical proof at a glance is not fundamentally different from a chess grandmaster's ability to evaluate a chess game at a glance. Along with the development of perceptual skill in chess there arises a kind of "aesthetic sensibility" for selecting good chess moves much like Poincaré's sensibility for good mathematics. It remains to be seen whether such observations can be worked into an explanation of Poincaré's intuitive powers that will satisfy both Miller and Simon. In the meantime, we can let Simon speak for himself in his book on scientific discovery (Langley, Simon, Bradshaw, & Zytkow, 1987).

We can now draw some important conclusions about imagery in physics, beginning with the fact that, in every branch of physics, the mathematics is accompanied by an informal system of imagery. On the one hand Miller has provided evidence that "imagery is a key ingredient in creative scientific thinking." On the other hand cognitive science provides evidence that imagery is equally important in the ordinary problem solving activities of physicists. There is no evidence for a qualitative difference in the imagery used in the two, cases, though the creation of a new imagery, such as Faraday's lines of force and Feynman's diagrams, may be a major scientific advance by itself.

The advance in our understanding of "imagery" will depend on how precisely we can define the concept. To that end, let us define an (objective) *image* as a pictorial (or diagrammatic) representation of information, and (objective) *imagery* as the manipulation of images to store, modify, or retrieve information. The "information content" of an image is determined by the rules of manipulation. The murky concept of *mental image* can then be regarded as a mental representation that is similar to an objective image, while *mental imagery* is mental manipulation of mental images. While the mental imagery of individual physicists may be highly variable, their objective imagery is fairly uniform, and "only the objective kind can be used to communicate ideas. Undoubtedly, the practice of objective imagery with all sorts of charts, graphs and diagrams helps a physicist develop his or her "physical intuition," which may well be a certain capability for mental imagery.

The imagery in physics is intimately related to mathematical representations of physical systems and processes. The imagery is related to the mathematics by a system of more-or-less definite rules. The rules may be explicitly formulated, as is the case for the "Feynman rules" relating Feynman diagrams to mathematical expressions. Often, however, the rules are tacit conventions that physicists employ "intuitively" without being aware of them. The lack of any clear explanation of the nature and function of physical imagery has been identified as a serious barrier to learning physics (Hestenes, 1987), forcing students to fumble about until they assimilate the imagery indirectly.

We can distinguish two different functions for imagery in physics. On the one hand imagery may perform a semantic function relating mathematics to experience. Einstein's *gedanken* experiments are of this kind. On the other hand imagery may perform a syntactic function exhibiting mathematical structure in a theory or model. Feynman diagrams, as well as other types of process diagram or system diagram, are of this kind. In general, diagrams perform information-compression (or chunking) and organizational functions. Chunking reduces information overload and facilitates the perception of wholes. And, as Einstein observed, images serve as "ordering elements for perception."

Imagery in physics is a promising domain for cognitive research. There is a rich lode of physical imagery that has never been mined systematically. Only a few prospectors like Miller and Simon have picked up samples. The payoff is likely to be greatest in education, leading to improvements in the design of images and in the teaching of imagery skills, thus enhancing creative powers at large. Here indeed, as Miller suggests, is a domain where historians and cognitive scientists can work together. But they had better enlist the help of some physicists.

REFERENCES

American Association of Physics Teachers (1963). *Special relativity theory: Selected reprints.* New York: American Institute of Physics.

Burtt, E. A. (1980). *The metaphysical foundations of modern science* (rev. ed.). New York: Humanities Press.

Chase, W., & Simon, H. (1981). Perception in chess. *Cognitive Psychology,* 4, 55-81.

Einstein, A. (1948). Dialog ueber Einwaende gegen die Relativitaetstheorie. *Naturnlissen-schaften,* 6, 697-702.

Fermi, E. (1950). *Nuclear physics.* Chicago, IL: University of Chicago Press.

Halloun, I., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics,* 53, 1056-1065.

Heitler, W. (1954). *The quantum theory of radiation* (3rd ed.). Oxford, U.K.: Clarendon Press.

Hestenes, D. (1979). Wherefore a science of teaching. *Physics Teacher,* **17,** 235-242.

Hestenes, D. (1987). A modeling theory of physics instruction. *American Journal of Physics,* 55, 440-454.

Kretschmann, E. (1917). Ueber den physikalischen Sinn der Relativitaetspostulate: A. Einsteins neue und seine urspruengliche Relativitaetstheorie. *Annalen der Physik,* 53, 575-614.

Langley, P., Simon, H., Bradshaw, G., & Zytkow, J. (1987). *Scientific discovery.* Cambridge, MA: MIT Press.

Moeller, C. (1972). *The theory of relativity* (2nd ed.). Oxford, U.K.: Clarendon Press.