The Mathematical Experience
by Philip Davis & Reuben Hersh
Birkhäuser, 1981

argument; Proof II exposes the "real" reason. In this way, the aesthetic component is related to the purer vision.

Further Readings. See Bibliography

S. A. Papert, [1978].

Nonanalytic Aspects of Mathematics

Conscious and Unconscious Mathematics

If we accept the common belief that the natural universe is governed by mathematical laws, then we understand that the universe and all within it are perpetually mathematizing—carrying out mathematical operations. If we are fanciful, we can think of each particle or each aggregate as the residence of a mathematical demon whose function it is to ride herd and say: "Mind the inverse square law. Mind the differential equations." Such a demon would also reside in human beings, for they, too, are constantly mathematizing without conscious thought or effort. They are mathematizing when they cross the street in fierce traffic, thereby solving mechanistic-probabilistic extremal problems of the utmost complexity. They are mathematizing when their bodies constantly react to transient conditions and seek regulatory equilibrium. A flower seed is mathematizing when it produces petals with six-fold symmetry.

Let us call the mathematizing that is inherent in the universe "unconscious" mathematics. Unconscious mathematizing goes on despite what anyone thinks; it cannot be prevented or shut off. It is natural, it is automatic. It does not require a brain or special computing devices. It requires no intellectual force or effort. In a sense, the flower or the planet is its own computer.

In opposition to unconscious mathematics we may distinguish "conscious" mathematics. This seems to be limited to humans and possibly to some of the higher animals. Conscious mathematics is what one normally thinks of as mathematics. It is acquired largely by special training. It seems to take place in the brain. One has special awareness of its going on or not going on. It is often tied to a symbolic and abstract language. It often is assisted by pencil and paper, mathematical instruments, or reference books.

But conscious mathematics does not always proceed through abstract symbols. It may operate through a "number sense" or a "space sense" or a "kinesthetic sense." For example, the problem "will this object fit into that box?" is answerable with high reliability on the basis of a mere glance. What lies behind these special senses is often not clear. Whether they represent stored experiences, analog-type solutions performed on the spot, inspired but partly random guesses, nonetheless the fact remains that this type of judgement can often be arrived at quickly and correctly. Although one is conscious of the problem, one is only partly conscious of the means by which the solution is brought about. Reflection after the fact often reveals a mixture of independent and overlapping operations. There is therefore no sharp dividing line between conscious and unconscious mathematizing.

Analog and Analytical Mathematics

It is convenient to divide conscious mathematics into two categories. The first, possibly more primitive, will be called "analog-experimental" or analog, for short. The second category will be called "analytic." Analog mathematizing is sometimes easy, can be accomplished rapidly, and may make use of none, or very few, of the abstract symbolic structures of "school" mathematics. It can be done to some extent by almost everyone who operates in a world of spatial relationships and everyday technology. Although sometimes it can be easy and almost effortless, sometimes it can be very difficult, as, for example, trying to understand the arrangement and relationship of the parts of a machine, or trying to get an intuitive feeling for a complex system. Results may be expressed not in words but in "understanding," "intuition," or "feeling."

In analytic mathematics, the symbolic material predominates. It is almost always hard to do. It is time consuming. It is fatiguing. It requires special training. It may require constant verification by the whole mathematical culture to assure reliability. Analytic mathematics is performed only by very few people. Analytic mathematics is elitist and self-critical. The practitioners of its higher manifestations form a "talentocracy." The great virtue of analytic mathematics arises from this, that while it may be impossible to verify another's intuitions, it is possible, though often difficult, to verify his proofs.

Insofar as the words analog and analytic are commonplace words that are used in many specific contexts in science, and since we intend them to have a special meaning in the course of this essay, we shall illustrate our intent with a number of examples. We begin with a very ancient problem which had a religious basis.

Problem: When is the time of the summer solstice, or of the new moon, or of some other important astronomical event?

Analog solution:

(i) Wait till it happens. Relay the happening to those concerned by messengers from the point of first detection.

(ii) Build some kind of physical device to detect important astronomical measures. Numerous astronomical "computers" are thought to have been used from prehistoric times in both the Old and New Worlds to detect the solstices and important lunar or stellar alignments, which are often of great importance for agricultural or religious reasons.

Analytic solution: Formulate a theory of astronomical periodicities and build it into a calendric structure.

Problem: How much liquid is in this beaker?

Analog solution: Pour the liquid into a graduated measure and read off the volume directly.

Analytic solution: Apply the formula for the volume of a conical frustrum. Measure the relevant linear dimensions and then compute.

Problem: What route should a bus take between downtown Providence and downtown Boston in order to maximize profits for the company?

Analog solution: Lay out a half dozen plausible routes. Collect time-cost-patronage data from the bus runs and adopt the maximizing solution.

Analytical solution: Make a model of the mileage, toll, and traffic conditions. Solve the model in closed form, if possible. If not, run it on a computer.

Analytical-existential solution: Demonstrate that on the basis of certain general assumptions, the calculus of variations assures us that a solution to the problem exists.

Problem: Given a function of two variables $f(x, y)$ defined over a square in the $x - y$ plane. It is desired to formulate a computer strategy which will give a plot of the contour lines of the function ($f(x, y) =$ constant).

Analytic solution: Starting at some point (x_0, y_0) compute $c = f(x_0, y_0)$. By means of inverse interpolation, find nearby points (x_1, y_1), (x_2, y_2), . . . for which $f(x_i, y_i) = c$. Connect these points. Iterate.

Analog-like solution: Place a fine grid over the square and think of the final picture as produced in a raster-scan fashion. Compute the function on the grid points and divide the range of values with say, 20 values: $v_1, \ldots, v_{20}$. Selecting a value v_i, draw in each small square either (a) nothing or (b) a straight line segment in the event that the four corner values are compatible with v_i. Iterate on i.

Contrasting Analog vs. Analytical Solutions

In some problems both analog and analytical solutions may be available. It may also happen that one is available and not the other, or that both are lacking. Neither type is to be preferred a priori over the other as regards accuracy or ease of performance. If both types of solutions are available, then the agreement of the two solutions is highly desirable. This may constitute the crucial experiment for a physical theory.

An attack on a problem is often a mixture of the two approaches. In the real world, an analytical solution, no matter how good, must always be fine tuned when a real system is to be modeled or constructed. Therefore, in engineering, the analytical solution is generally taken as a point of departure, and, we hope, a good first approximation.

The analog solution appears to be closer to the unconscious mathematizing that goes on in the universe. Analog solutions probably predominate in the world of technology —but this is a pure conjecture.

The Hierarchy of Intellectual Values

When it comes however to the intellectual value that is set on these two modes, it is clear what the ordering is. Although an analog solution may be clever, based on sophisticated and subtle instrumentation, it does not carry the accolade of the purely intellectual solution.

The intellect looks after its own. What is consciously more difficult is the more praiseworthy. The level of intellectual acclaim is proportional to the apparent complexity of the abstract symbolization. Up to a point, of course; for the house built by intellect may crash to the ground when confronted by experimental reality. Scientific education is often directed not to the solution of specific problems, but toward the carrying on of a discourse at the highest possible intellectual level.

The hierarchy we have just described is that of the practicing mathematician. From the point of view of an engineer, say, an analytical solution is of little interest unless it leads to a functioning device, which corresponds to an analog solution to a problem. A well-designed, highly developed device can show the economy of means and the elegance of thought that characterizes the best science and mathematics, but this elegance is seldom recognized by theoretically oriented scientists. The ingenuity and artistry that go to make a functioning airplane or a reliable, efficient computer are seldom appreciated until one tries to do something along these lines oneself.

Similar Hierarchies Exist in the Nonscientific World

Nor is this tendency toward stressing the intellect limited to science. It occurs, for example, in the art world. At the very lowest level is the commercial artist. Somewhat higher is the artist who paints portraits on demand. At the highest level stands the "fine artist" who is supposed to respond to the abstract and unfettered promptings of the intellect and spirit. The work of art is often accompanied by an explication de texte whose abstraction may rival the deepest productions of mathematics.

Mathematical Proof and its Hierarchy of Values

The definition-theorem-proof approach to mathematics has become almost the sole paradigm of mathematical exposition and advanced instruction. Of course, this is not the way mathematics is created, propagated, or even understood. The logical analysis of mathematics, which reduces a proof to an (in principle) mechanizable procedure, is a hypothetical possibility, which is never realized in full. Mathematics is a human activity, and the formal-logical account of mathematics is only a fiction; mathematics itself is to be found in the actual practice of mathematicians.

An interesting phenomenon should be noted in connection with difficulties of proof comprehension. A mathematical theorem is called "deep" if its proof is difficult. Some of the elements that contribute to depth are nonintuitiveness of statement or of argument, novelty of ideas, complexity or length of proof-material measured from some origin which itself is not deep. The opposite of deep is "trivial" and this word is often used in the sense of a put-down. However, it does not follow that what is trivial is uninteresting, unuseful, or unimportant.

Now despite this hierarchical ordering, what is deep is in a sense undesirable, for there is a constant effort towards simplification, towards the finding of alternative ways of looking at the matter which trivialize what is deep. We all feel better when we have moved from the analytic toward the analog portion of the experiential spectrum.

Cognitive Style

An obvious statement about human thought is that people vary dramatically in what might be called their "cognitive style," that is, their primary mode of thinking.

This was well known to nineteenth-century psychologists. Galton, in 1880, asked a wide range of people to "describe the image in their mind's eye of their breakfast table on a given morning." He found that some subjects could form vivid and precise pictures while others could form only blurry images or, in some cases, no image at all. William James reported that people varied greatly in the sense modality they primarily used to think in, most people being auditory or visual. There was a smaller number, however, who were powerfully influenced by the sense of touch or of kinesthesis (movement), even in what is usually called abstract thought.

Such a wide range of ways of thinking should cause no problems. Indeed, we might regard with pleasure the diversity of methods of thinking about the world that our species shows, and value all of them highly as valid ways of approaching problems.

Unfortunately, tolerance is a rare virtue, and a common response to different ways of thinking is to deny, first, that they are possible, and, second, that they are valuable.

William James remarks, "A person whose visual imagination is strong finds it hard to understand how those who are without the faculty can think at all."

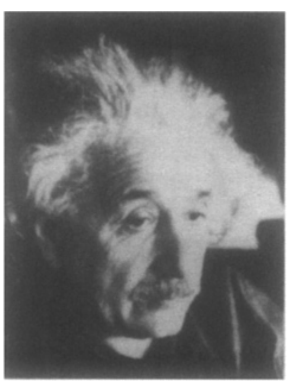
Willard V. O. Quine 1908–

Conversely, some of those who think mostly in words are literally incapable of imagining nonlinguistic thought. W. V. O. Quine says that ". . . memories mostly are traces not of past sensation but of past conceptualizations or verbalizations." Max Muller wrote, "How do we know that there is a sky and that it is blue? Should we know of a sky if we had no name for it?" He then argues that thought without language is impossible. Abelard said, "Language is generated by the intellect and generates intellect." The Chandogya Upanishad states, "The essence of man is speech."

The Gospel of John starts, "In the beginning was the word. . . ."

However, many have held opposing views. Aristotle said we often think and remember with images. Bishop Berkeley held that words are an impediment to thought. Many philosophers and theologians view concepts and words as dangerously misleading "word play." The Lankavatara Sutra is typical, "Disciples should be on their guard against the seductions of words and sentences and their illusive meanings, for by them, the ignorant and dullwitted become entangled and helpless as an elephant floundering around in deep mud. Words and sentences . . . cannot express highest reality. . . . The ignorant and simple minded declare that meaning is not otherwise than words, that as words are, so is meaning. . . . Truth is beyond letters and words and books." The Tao Te Ching (LXXXI) says, "True words are not fine sounding; fine-sounding words are not true. The good man does not prove by argument and he who proves by argument is not good. . . ." A Biblical quotation in this tradition says, "The letter killeth, but the spirit giveth life. . . ."

Cognitive Style in Mathematics



Albert Einstein
1879–1955

In his book, Hadamard tried to find out how famous mathematicians and scientists actually thought while doing their work. Of those he contacted in an informal survey, he wrote "Practically all of them . . . avoid not only the use of mental words, but also . . . the mental use of algebraic or precise signs . . . they use vague images." (p. 84) and ". . . the mental pictures of the mathematicians whose answers I have received are most frequently visual, but they may also be of another kind—for example kinetic." (p. 85)

Albert Einstein wrote to Hadamard that "the words or the language, as they are written or spoken, do not seem to play any role in my mechanism of thought. . . . The physical entities which seem to serve as elements in thought are certain signs and more or less clear images which can be 'voluntarily' reproduced and combined. . . . The above mentioned elements are, in my case, of visual and some of muscular type. Conventional words or other signs have to be sought for laboriously only in a secondary stage . . ." (p. 142) Several recent studies on the way in which nonmathematical adults perform simple arithmetic seem to suggest the same is true for nonmathematicians as well.

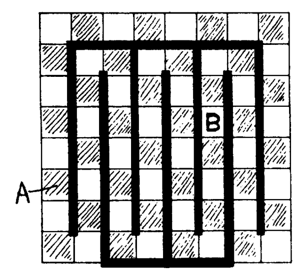
An Example of Cognitive Style in Combinatorial Geometry

We have already described and contrasted analog solutions and analytical solutions. Insofar as mathematical discovery may have large components of one or the other, we are dealing here with differences in cognitive style.

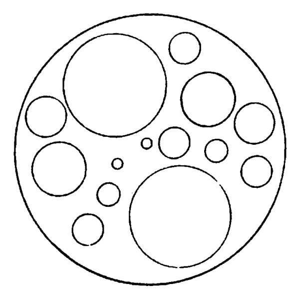
Here is a striking example where an analytic proof might be very difficult while an analog-like proof makes the whole business transparent.

Gomory's Theorem. Remove one white and one black square from an ordinary checkerboard. The reduced board can always be covered with 31 dominos of size 2 × 1.

Analog Proof. Convert the checkerboard into a labyrinth as in the accompanying figure. No matter which black square "A" and which white square "B" are deleted, the board can be covered by threading through the labyrinth with a caterpillar tractor chain of dominos which break off at "A" and "B." (See Honsberger, p. 66.)



The imagery of a moving caterpillar tread is sufficient to enable one to grasp the solution at a glance. Note the powerful kinesthetic and action-oriented mode of proof. It would be difficult for the normal (i.e., somewhat visual) reader to work through this proof and not feel a sense of movement.



We are not aware of an analytical solution to the problem. Of course one could tighten up the above solution by counting blacks and whites to provide a more formal proof.

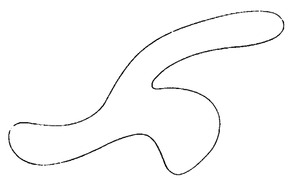
Here is a geometrical problem where both types of solution are available.

Theorem. *It is impossible to fill up a circle C with a finite number of nonoverlapping smaller circles contained in C.*

Analog Solution. This is visually obvious.

Analytical Solution. For a neat proof based on notions of linear independence, see Davis, 1965. The analog solution is so apparent that to insist on more is a piece of mathematical pedantry.

This leads us to the notorious



Jordan Curve Theorem. A simple closed curve in the plane separates the plane into two regions, one finite and one infinite.

Analog Solution. This is visually obvious.

Analytical Solution. Very difficult, the difficulty deriving from the fact that an excessive degree of analytical generality has been introduced into the problem.

Mathematical Imagery

Hadamard described the semiconscious stream of thought which may accompany the process of conscious mathematizing. This kind of thing surely exists although it is very difficult to describe and document.

A few more words along this line describing my own experiences might therefore be appropriate.

The semiconscious stream of thought—which might be referred to as mathematical imagery—does not seem to relate directly to the analytical work attempted. It feels to be more analog, almost visual, sometimes even musical. It accompanies and occasionally helps the dominant stream of thought. It frequently seems irrelevant, a mere hovering background presence.

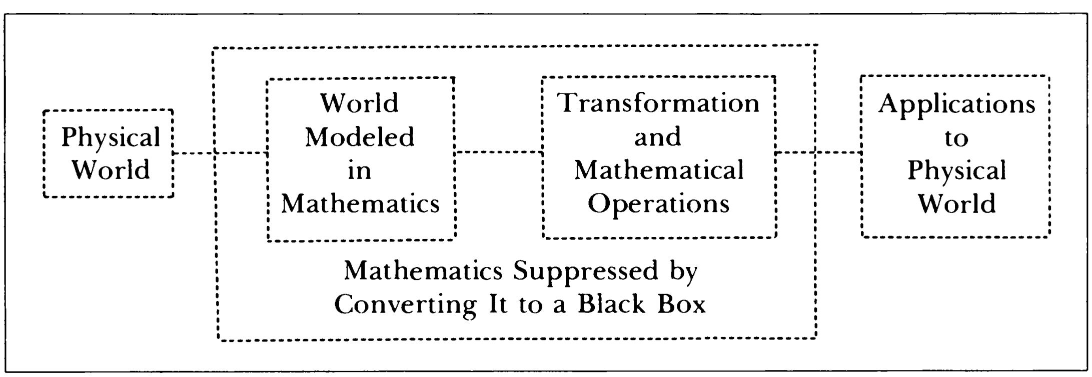
Some years ago I spent considerable time working in the theory of functions of a complex variable. This theory has a considerable geometric underlay. In fact, it can be developed independently from a geometric (Riemannian) or from an analytic (Weierstrassian) point of view. The geometrical illustrations in textbooks often feature spheres, maps, surfaces of an unusual kind, configurations with circles, overlapping chains of circles, etc. As I was working along with the analytic material, I found it was accompanied by the recollection or the mixed debris of dozens of pictures of this type that I had seen in various books, together with inchoate but repetitious nonmathematical thought and musical themes.

I worked out, more or less, a body of material which I set down in abbreviated form. Something then came up in my calendar which prevented me from pursuing this material for several years. I hardly looked at it in the interval. At the end of this period, time again became available, and I decided to go back to the material and see whether I could work it into a book.

At the beginning I was completely cold. It required several weeks of work and review to warm up the material. After the time I found to my surprise that what appeared to be the original mathematical imagery and the melody returned, and I pursued the task to a successful completion.

The Proper Goal of Mathematical Applications is for the Mathematics to Become Automated

Analytical mathematics is hard to do and is inaccurately performed if done quickly. We do not expect that an astronaut landing on the moon will be laying out his actions as the result of computing with a table of logarithms and trigonometric functions in real time. Ideally, as in unmanned flight, the whole system might be automated. Although there is a considerable mathematical underlay to landing manoeuvres, we expect that the astronaut will



respond to instrument or computer readouts or verbal instructions which are surrogate readouts. The analytical mathematics must be suppressed or bypassed and replaced by analog mathematics.

One sees this over and over again in applied mathematics. The more complete and successful is an application, the more automatic, programmed, rote, it must become.

An Example From Computer Graphics

A striking example of the suppression of the mathematical underlay occurred in the last ten to fifteen years in the area of computer art and animation.

Computer art can be traced back to mechanical instruments that produced cyclic motions of various sorts, hypocycloids, Lissajou's figures, "spirographic" figures, etc. These were easily produced on oscillators or analog circuitry. The idea was to draw figures which were visually pleasing or exciting. There was a constant identification of the figure with the mathematical equation.

In the first decade or so of computer art, the mathematical presence was quite noticeable. In addition to having artistic sensibilities, practioners had to know computer programming, graphical programming, a certain amount of basic mathematics such as analytic geometry, elementary transformations, and interpolation schemes. Gradually, higher and higher level languages were written for computer art. The mode of operation became less analytic, more linguistic, and more analog. As this occurred the mathematical substructure was either built in, suppressed, or bypassed.

An excellent instance of this suppression is the PAINT program developed at the University of Utah and New York Institute of Technology. As a response to a desire to do commercial animation by computer, a very high level language was developed which could be learned easily and used by commercial animators without a knowledge of mathematics. Working in color, the artist is able to select a palette and create brushes of various widths and spatter characteristics. He creates shapes working with a stylus on a computer sketch pad. Numerous "menu items" allow him to fill with color, to exchange colors continuously in real time, to replicate, to transform (zoom), to phantom, to animate, via linear and nonlinear interpolation.

Thus the artist inputs to the computer his own muscle movements of wrist and arm and shoulder as well as a set of "menu options." Since these options are also controlled by the stylus, the whole process proceeds in imitation of the manner of painting in conventional media.

The Degradation of the Geometric Consciousness

It has often been remarked over the past century and a half that there has been a steady and progressive degradation of the geometric and kinesthetic elements of mathematical instruction and research. During this period the formal, the symbolic, the verbal, and the analytic elements have prospered greatly.

What are some of the reasons for this decline? A number of explanations come to mind:

(1) The tremendous impact of Descartes' *La Géométrie*, wherein geometry was reduced to algebra.

(2) The impact in the late nineteenth century of Felix Klein's program of unifying geometries by group theory.

(3) The collapse, in the early nineteenth century, of the view derived largely from limited sense experience that the geometry of Euclid has a priori truth for the universe, that it is *the* model for physical space.

(4) The incompleteness of the logical structure of classical Euclidean geometry as discovered in the nineteenth century and as corrected by Hilbert and others.

(5) The limitations of two or three physical dimensions which form the natural backdrop for visual geometry.

(6) The discovery of non-Euclidean geometries. This is related to the limitations of the visual ground field over which visual geometry is built, as opposed to the great generality that is possible when geometry is algebraized and abstracted (non-Euclidean geometries, complex geometries, finite geometries, linear algebra, metric spaces, etc.).

(7) The limitations of the eye in its perception of mathematical "truths" (continuous, nondifferentiable functions; optical illusions; suggestive, but misleading special cases).

An excellent exposition of the counter-intuitive nature of analytical mathematics when it attempts to extend the visual field can be found in the article by Hahn in volume III of *The World of Mathematics.* It has become traditional to regard these "pathological examples" as pointing to failures of the visual intuition. *But they can equally well be interpreted as examples of the inadequacies of the analytical modeling of the visual process.*

Right Hemisphere and Left Hemisphere

There is an intriguing, but speculative similarity between the two approaches to mathematics we have described and current work on the functions of the two cerebral hemispheres. Although this work is still in its beginnings, it seems clear that the right and left hemispheres are specialized to do somewhat different tasks. (For a review of this rapidly growing field, see part one of Schmitt and Worden, and for an informal discussion, see Gardner, *The Shattered Mind.*)

It has been known for over a hundred years that in virtually all right-handed humans and in about half of the left-handed, the parts of the brain associated with speech are primarily located in the left cerebral hemisphere. This appears to be an innate biological specialization, and a slight anatomical asymmetry has been shown to exist between the hemispheres in both newborn infants and adults. Damage to certain areas of the left hemisphere will cause characteristic types of difficulties with speech, while damage to the right hemisphere in the same location will not. To oversimplify a complex issue, the left hemisphere in most humans is primarily concerned with language-based behavior and with the cognitive skills we might crudely characterize as analytical or logical. It has become apparent recently that the right hemisphere is far superior to the left in most visual and spatial abilities, discriminations by touch, and in some nonverbal aspects of hearing, for example, music.

A great deal of information about hemispheric specialization has come from careful study of a small number of neurosurgical patients who had to have the two hemispheres disconnected as a last resort against life-threatening epilepsy. Sperry has summarized many of the results of research on these patients:

> Repeated examination during the past 10 years has consistently confirmed the strong lateralization and dominance for speech, writing, and calculation in the disconnected left hemisphere in these right handed patients. . . . Though predominantly mute . . . the minor hemisphere is nevertheless clearly the superior cerebral member for certain types of tasks. . . . Largely they involve the apprehension and processing of spatial patterns, relations, and transformations. They seem to be holistic and unitary rather than analytic and fragmentary . . . and to involve concrete perceptual insight rather than abstract, symbolic, sequential reasoning. (p. 11)

We should always remember that the two hemispheres combine to give a whole brain. Even in speech, the left hemisphere function, the right hemisphere plays an important role, and the hemispheres work in harmony in a normal individual.

The melodic aspects of speech—rhythm, pitch, and intonation—seem related to the right hemisphere. Gardner provides an apt description of other functions:

> With individuals who have disease of the right hemisphere, the abilities to express oneself in language and to understand . . . others are deceptively normal . . . [however] these patients are strangely cut off from all but the verbal messages of others. . . . They are reminiscent of language machines. . . . appreciative of neither the subtle nuances or non-linguistic contexts in which the message was issued . . ." (p. 434)

Gardner then goes on, in a way somewhat unflattering to the public image of mathematicians:

> Here the patient (with right hemisphere damage) exemplifies the behavior . . . associated with the brilliant young mathematician or computer scientist. This highly rational individual is ever alert to an inconsistency in what is being said, always seeking to formulate ideas in the most airtight way; but in neither case does he display any humor about

> his own situation, nor . . . the many subtle intuitive interpersonal facets which form so central a part of human intercourse. One feels rather that the answers are being typed out at high speed on computer printout paper. (p. 435)

The anecdotes we have given earlier, and our own experiences, indicate that mathematics makes use of the talents that are found in both hemispheres, rather than being restricted to the linguistic, analytic specialties of the left hemishere. The nonverbal, spatial, and holistic aspects of thought are prominent in what most good mathematicians actually do though perhaps not so much in what they say they do.

It is a reasonable conclusion that a mathematical culture that specifically downgrades the spatial, visual, kinesthetic, and nonverbal aspects of thought does not fully use all the capacities of the brain.

The de-emphasis of the analog elements of mathematics represents the closing off of one channel of mathematical consciousness and experience. Surely, it would be better to develop and use all the special talents and abilities of our brains, rather than to suppress some by education and professional prejudice. We suggest that in mathematics it would be better for the contributions of the two halves of the brain to cooperate, complement, and enhance each other, rather than for them to conflict and interfere.

Further Readings. See Bibliography

T. Banchoff and C. Strauss; P. J. Davis [1974]; P. J. Davis and J. Anderson; H. Gardner; I. Goldstein and S. Papert; J. Hadamard; R. Honsberger; W. James [1962]; M. Kline [1970]; K. Knowlton; Wikuyk; E. Michener; R. S. Moyer and T. K. Landauer [1967], [1973]; S. Papert [1971]; M. Polányi [1960]; R. D. Resch; F. Restle; F. O. Schmitt and F. G. Worden; R. W. Sperry; C. M. Strauss; R. Thom [1971].