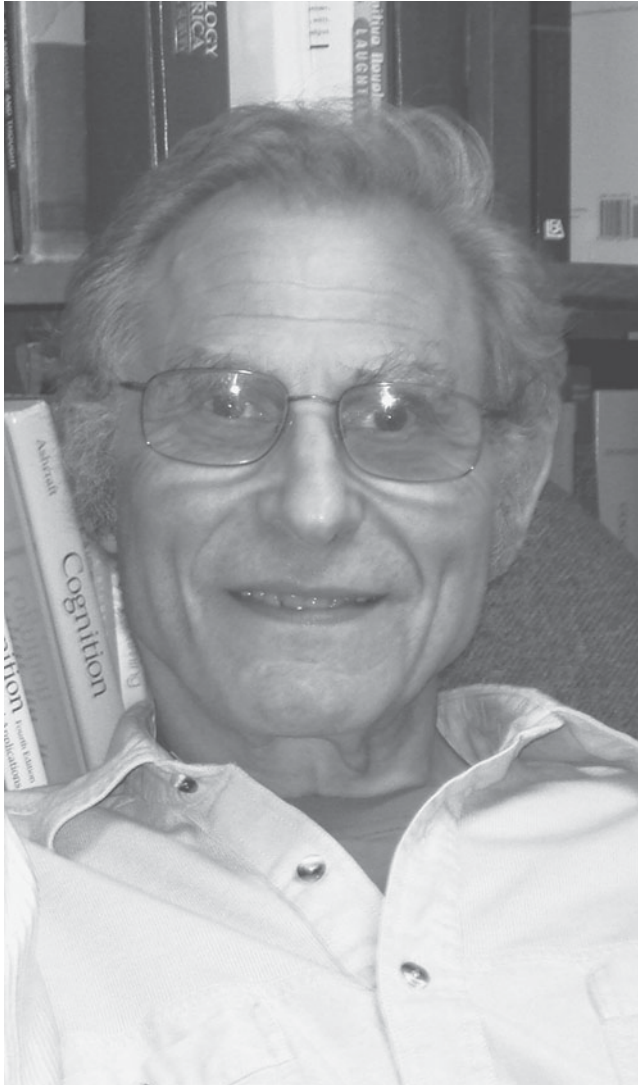


Michael Posner



Michael I. Posner

BORN:

Cincinnati, Ohio
September 12, 1936

EDUCATION:

University of Washington, B.S. (1957)
University of Washington, M.S. (1959)
University of Michigan, Ph.D. (1962)

APPOINTMENTS:

Assistant Professor of Psychology, University of Wisconsin (1962–1965)
Associate Professor of Psychology, University of Oregon (1965–1968)
Professor of Psychology, University of Oregon (1968–EMERITUS 2000–present)
Professor of Neuropsychology and Psychology, Department of Neurology Neurosurgery,
Washington University, St. Louis (1985–1989)
Director, Institute of Cognitive and Decision Sciences, University of Oregon (1989–1995)
Head, Department of Psychology, University of Oregon (1995–1998)
Professor of Psychology in Psychiatry, Weill Medical College of Cornell University, New York
Founding Director, Sackler Institute (1998–2002)
Adjunct Professor, Weill Medical College (2002)

HONORS AND AWARDS (SELECTED):

Paul M. Fitts Award, Human Factors Society (for training in Human Factors, 1967); Ersted Award
for distinguished teaching, University of Oregon (1975); John Simon Guggenheim Fellow
(1979–1980)
A.P.A. Distinguished Scientific Contribution Award (1980)
National Academy of Sciences, U.S.A. (1981)
American Academy of Arts and Sciences (1986)
Howard Crosby Warren Medal, Society of Experimental Psychologists (1988)
Howard Vollum Award for exceptional achievement of a member of the scientific and technical
community of the Pacific Northwest (1989)
Scientist of the Year Award, Oregon Academy of Sciences (1995)
Charles A. Dana Award for pioneering research in medicine (neuroscience) (1997)
Honorary Doctoral Degrees: Universities of Padova (1998), Granada (1999), Nottingham (2002),
Paris (2002), Brussels (2007), Ben Gurion (2007), Michigan (2009), and Buenos Aires (2010)
Karl Lashley Award, American Philosophical Society (joint with M. E. Raichle) (1998)
Pasarow Foundation Award in Medical Research (Neuropsychiatry) with M. E. Raichle (2000)
First Grawemeyer Award for Psychology with M. E. Raichle and S. E. Petersen (2001)
Fyssen Foundation International Prize for Studies of Human Consciousness (2004)
Patricia Goldman Rakic Award NARSAD (with M. E. Raichle) (2004)
First Mattie Dogan Prize in Psychology (2008)
National Medal of Science (2008)

Michael I. Posner's early work involved the measurement of mental operations by use of reaction time and other chronometric measures. In 1979 he began studies of patients to link mental operations to brain areas in the study of attention. To test hypotheses arising from this work, he developed with Marcus Raichle studies imaging the human brain during cognitive tasks. He has worked on the anatomy, circuitry, development, and genetics of three attentional networks underlying maintaining alertness, orienting to sensory events, and voluntary control of thoughts and ideas. His methods for measuring these networks have been applied to a wide range of neurological, psychiatric, and developmental disorders and to normal development and school performance. At the time of this volume his research involves a longitudinal study of children prior to school designed to understand the interaction of specific experience and genes in shaping attention and self-regulation and studies of methods for training attention networks and changing the brain state in ways that might also alter attention.

Michael I. Posner

A Life in Psychology

I was fortunate to be around during extraordinary times for the study of psychology. Neuroimaging made the human brain, which is obviously central to understanding the mind and behavior, available for scientific analysis. This biography tries to trace some of the background for my involvement in these events. I believe that a better understanding of the human mind requires both measurement of mental events and mapping of those events to brain networks. It has been my goal over many years to contribute to that effort.

Personal Story

I was born late in the Great Depression on September 12, 1936, in Cincinnati, Ohio, but at two months of age I travelled with my family to San Bernardino, California. My father had a successful law practice in Cincinnati, but for health reasons he undertook the move to California. Because California had no reciprocity in law degrees and because of the difficult economic times during the depression years he never again practiced law. During the war years he first worked in defense-related industry, including Cal Ship and later social work with Jewish servicemen and their families. My mother was most often home with us in these years, but she also was a very fine shoe salesperson, a calling she pursued mostly on weekends.

The most influential people in my early development were my father, whose extraordinary dedication to helping others was always a powerful example in my life, and my brother Jerry, whose guidance has continued over my whole career. He was a brilliant student who decided to become a physician at age 2, received his M.D. in his early 20s and became a very renowned neurologist and founder of the Department of Neuro-Oncology at Memorial Sloan Kettering Cancer Center in New York. After my graduation in physics he suggested graduate work in biological science, which eventually led to my choice of psychology.

In 1979 I went to New York and learned from Jerry Posner about how to study neurological patients. In my 3 years in the Neurology and Neurosurgery Departments in St. Louis at the start of cognitive imaging, being Jerry's brother helped to overcome my connection to psychology and cognition, both regarded with suspicion in that department. My role as a mystic in St. Louis is described somewhat in the section on the neural basis of

mental operations. Jerry's reputation also aided me during my 3 years in Psychiatry at Weill Medical College, which is described in the section on attention.

Measuring Mental Operations

In 1963 I was a young Assistant Professor at the University of Wisconsin in Madison. I had been appointed in 1962 and set up a small laboratory in the old Journalism building, which was just behind 600 N. Park where Psychology was located. I thought it important to do something in this laboratory. It was the department of Harry Harlow, David A. Grant, and Wilf Brogden. They were important in psychology and thought that anyone in their department should be as well. It was not the easiest department to join!

Inspired by Broadbent's 1958 book, I planned to study attention. I had set out to study the time to switch attention from one modality to another. This was the era before computers, so in the experiment I used a gift from my mother-in-law of a Wallensak tape deck together with a new display device consisting of small photos of the digits 1-9 and another with upper- and lowercase letters. When a light was switched on, the digit or letter was projected so that it could be seen on the front screen. This in-line display was switched from a beep on one channel of the recorder. In this way I could expose the subject to an auditory stimulus (e.g., letter or digit) and a visual digit or letter. The subject's task was to say whether the two successive items were the same or different. My goal was first to determine how much longer it took to deal with matching cross-modal events than events in the same modality. Then to determine the switching time I planned to delay the second item until the difference went away and the length of the delay needed would be a measure of the switching time needed.

If you were to have read my papers that appeared a few years later based on these findings, you would never have guessed the purpose of my study. In the course of the studies I found out something quite different and, from my perspective, more important. It does take longer to match an auditory digit following a visual one than to match two identical auditory or two identical visual digits, just as I had predicted. However, when the task was changed from matching the identical digits to one of determining whether digits were odd or even, the time for switching went away. It no longer took time to switch between modalities. Instead I found that whenever the two items were not physically identical there was an increased time. If, for example, I had the subjects indicate whether two letters had the same name and the pair were upper- and lowercase (e.g., Aa versus AA or aa), I found the same 80 millisecond increase that occurred for matching a visual and auditory A. Why was Aa longer than AA? From my perspective steeped in the learning theory of the 1960s, this was a genuine puzzle. We have spent a lifetime learning that Aa were the same letter, and even longer learning that a visual

and auditory 3 were both the same digit. If our mental processes were based on learned associations as it said in all my learning textbooks, why were physically identical items at such an advantage?

Perhaps matching a capital and small letter was longer than pairs with the same name because the identical letter pairs were identical in two ways. They both had the same name and they were the same shape. If this were true, I reasoned that pairs like AB, which were different in both form and name, would take the same time irrespective of whether the instruction were to match based on physical identity or on name. However, I found in another experiment that it took 80 milliseconds longer to match AB when the matching instruction was based on physical identity (i.e., are the two letters exactly identical?) than when it was based on name (i.e., do the two letters have the same name?). Since AB were always just as similar, I thought this meant that some internal process involved in determining whether a nonidentical pair has the same name had to be performed before one could decide AB did not require a yes to the name match instruction. I was measuring a covert mental process that delayed making the no response to letter pairs like AB. I loved this idea.

Of course I knew that 100 years ago the great Dutch physiologist Donders had first measured the time for the covert mental processes of recognition and choice by using the subtractive method. I also knew that the subtractive method has been criticized by a number of psychologists as requiring that all processes remained the same except the one being measured, but since the instruction changed from one experiment to another there was no way to be sure this was true. In fact, the leading historian of psychology, Boring, had dismissed Donders' method as useless, in his history of psychology that all students of my generation had been required to read. However, in my study the 80 millisecond difference between AA and Aa could be obtained in exactly the same experiment with the same instruction. Thus, I reasoned it was not as subject to the criticism made of Donders. At the time I could not have guessed that 30 years later I would still be debating this issue with a generation of neuroscientists who had never been exposed either to Donders or Boring, but were using subtraction in neuroimaging experiments.

My paper describing these findings appeared in the *Psychological Review* in 1967. By then I had left Madison, Wisconsin, to take up a position at the University of Oregon in Eugene, which was to be my lifetime home. In my work I felt that the covert mental process that I had trapped might be a good empirical method for measuring the kinds of internal computations that were discussed in a new book by Miller, Galanter, and Pribram that took the ideas of Herbert Simon and applied them to the questions of experimental psychology. At Oregon there were already a number of psychologists closely related to the topics on which I was working. Fred Attneave had written a great book on how information theory might relate to psychology, and Ray Hyman was one of the discoverers of the law that related the time taken to

perform a task to the amount of information contained in the task. The trips to Oregon were full of discussion about how to shape the new information processing psychology that was emerging from many laboratories.

Setting up a new laboratory is always difficult. In Eugene it was more of a problem because the Psychology department had almost no space. It had a small animal laboratory in which Jim McGaugh had done many learning studies. Jim left the year I came and in fact I was his replacement and was to teach courses in learning which he had formerly taught. Dan Kimble, who was already part of the faculty, was conducting research on the hippocampus. Fred Attneave had a small room in Condon Hall where psychology offices were located, and Ray Hyman had space in Susan Campbell a few buildings away, but there was no substantial laboratory for human studies. Indeed in human work most of the effort was on the development of theory, and experiments played a smaller role. We were given a house on the far end of campus, which had been a fourplex. I had start-up funds of \$3500 to equip a laboratory. In these days where some new faculty receive up to half a million dollars, that might seem a bit small, even corrected for inflation. I traveled with Dick Littman (then department head) to Salem, where the state warehouse had furniture, much of it made in the prison, which was sold cheaply to qualified state agencies. I remember that in addition to new visual displays I also bought a Stowe memory drum, a tachistoscope, and slide projector. Within a few years this laboratory was to house four faculty, a dozen students, and more than \$100,000 worth of computer equipment, but that was all in the future.

The paper on letter matching, written with graduate student Ron Mitchell, was published in 1967. I felt that we had made a very considerable contribution in showing how to improve the old subtractive method. However, a much greater contribution to this problem was to come from another source the very next year. In appreciation of the 100th anniversary of Donders' paper, the new group called Attention and Performance, which had been started by Andries Sanders in the summer of 1966, held a meeting in Eindhoven, the Netherlands. On the last day of that meeting Saul Sternberg presented his paper on the additive factors method for studying how to infer serial stages from reaction times. Everyone at the meeting knew immediately that this was truly a great contribution. It showed how Donders' insights could be extended to any serial task by examining additivity and interactions between independent variables thought to influence the same or different internal stages. My own reaction was somewhat mixed. I knew this was a great contribution to the field of psychology. I also knew it would eclipse whatever my own studies based on the more limited matching method had produced. While there was clearly a bit of jealousy for what Sternberg had achieved, a feeling I was to have many times over the years, I was also happy that at last we had a strong, if limited, method to explore the internal workings of the mind.

What exactly were the mental operations involved in matching an upper- and lowercase A? I first imagined that the person determined whether the two letters were identical in physical form and, if the answer was no, then tested whether they had the same name. This would be a strictly serial model. I rejected this idea because matching at the physical level would take time and, if all the physical matching stimuli were left out so that subjects would have to name all name pairs, no reduction in RT resulted. It seemed like the two processes took place in parallel. If so, the ending times for each process had to be variable because we could extend the time for physical matches by making the color different between two matching letters and not influence the name-level matches. Having other letter names in memory increased the time for the name match but not the physical match. We seemed to have two relatively independent and parallel processes.

In 1967 when these experiments were being carried out, Richard Atkinson and Richard Shiffrin published a joint paper on memory for letters and other materials. In this very famous model there was an encoding process that converted the visual letters to a verbal code that was stored. There was no idea that the physical form of the letter could also be stored. It was a strictly serial model of the type that was then common in psychology and in physiology. Additive factors in psychology built upon the serial processing ideas that were the basis of digital computers at that time, and this appeared to fit with the kind of hierarchy of simple, complex, and hypercomplex cells that had been found in the visual system by Hubel and Wiesel. Parallel models were thought of as too complex and as violations of the scientific strictures for simplicity. The massively parallel systems at the basis of connectionism were 15 years in the future.

Steve Keele and I thought we could study the possibility of parallel coding in memory if we simply put in one letter and then delayed the time before the second letter was presented. We found that at least for several seconds after the first letter disappeared, the trace of the physical form of the first letter must have been present because physical matches were still faster than name matches. Moreover, we could extend the life of the physical form by making name matches more difficult, so that even when we measured no advantage in the standard experiment that did not mean that the physical form was no longer stored. We published the first of these results in a paper in *Science*, which I believe was the first evidence against the serial coding model in memory.

Steve and I realized that we could test the imagery ability of subjects by making the first letter auditory and seeing how they might respond to upper- and lowercase matches. When instructed to code an auditory stimulus as an uppercase A, a delay of .5 sec was sufficient for the match to be identical to the time for matching against a physically presented uppercase A. Moreover, when this occurred the time for a lowercase A was the same length as for letters not presented on that trial. Based on our letter match results,

Barbara Tversky taught subjects names to correspond with simple line drawings of face-like stimuli with varying numbers of features in common. Later, Miriam Klein (now Rogers) varied the physical similarity or the similarity of the names. The findings suggested that when given the name of a face name, subjects could retrieve the physical face in about .5 of a second. At that time most subjects claimed to have a visual image. However, earlier when most people denied having an image, the match still depended upon the number of visual features in common. I distinguished between a visual code in memory, which clearly influenced the time to match stimuli, and a visual image that was available to consciousness.

At the time we published a monograph illustrating the dual nature of codes and the ability to maintain a visual code in memory and to create an image in .5 second (Posner, Boies, Eichelman, & Taylor, 1968) it was still thought that visual images were somewhat mystical. How could something as subjective as a visual image be measured and thus meet the criterion of behaviorism for objective evidence? I believed then that our evidence putting auditory letters into concrete visual form provided such evidence. However, a year or two after our paper, Roger Shepard published a series of papers showing that one could create a visual representation of a letter and rotate it at will to any angle. Even more beautiful was the finding that, once having created an image at a particular angle, one could match a rotated letter faster than even an upright one. A lifetime of learning to deal with upright letters was less important for matching than a created representation at the instructed angle. This was beautiful evidence for imagery and it was to be shortly supplemented and massively documented by the studies of Steve Kosslyn showing in detail the properties of visual images. Moreover, these important findings set the stage for the results Kosslyn was to obtain with positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) showing how even primary visual cortex could be influenced by images.

I must admit a kind of quiet pride at all of these results on imagery. I think it is one of the great cumulative success stories of psychology in the current era. At times I felt a bit upset when a student would ask me whether I knew about the discoveries on images of Shepard or Kosslyn, but I realize that one needs to remember the very best evidence that was obtained on a topic, which is not necessarily the same as the first evidence. Even when I lecture on imagery I refer mainly to mental rotation or to Kosslyn's many demonstrations. I still feel that the distinction between code and image is important for psychology, even more these days when neuroimaging studies show late activation within sensory areas in tasks where no conscious image occurs.

Exemplars and Prototypes

At the time of my dissertation studies, information theory was at the height of its popularity in psychology. It had been very successful in summarizing

how the amount of information transmitted by a stimulus could provide a prediction of the time required to respond. This finding suggested that human mental activity could be measured in terms of channel capacity or the bits per second that could be transmitted by the system. It thus extended the subtractive method by allowing one to deal with the probability of a stimulus and error along with the number of stimuli by a single measure that could be correlated with reaction time. At last Psychology had a law: reaction time was proportional to information transmitted (Hick-Hyman law).

My dissertation adviser, Paul Fitts, used information theory to extend the law to movement. When the extent and accuracy of a movement was related to the amount of information that the movement conveyed, one could predict the time to move to the target quite accurately (Fitts law). The two laws together, and knowledge of a person's speed, could allow one to predict when a person would leave and reach a wide variety of targets. These findings were the centerpiece of the book, *Human Performance*, I was to finish for Fitts after his very untimely death.

It was difficult to describe the pleasure of being able to have two laws based upon the same underlying theory. More than just the new laws, information theory provided a critique for pure behaviorism. It was not just reinforced practice that influenced speed of processing. Instead what might have occurred but did not, that is, the alphabet of possible stimuli, was also important for prediction. This suggested that some intrinsic structure limits the human capacity to process information. It was our goal to understand these limits, and this was the basis for my strong concentration on attention.

As soon as one left the experimental tasks that Fitts had used, there were clearly problems. Consider adding two numbers together. Here was a mental operation that did not preserve the information in the stimulus, but of course it required time to calculate the response. Mental arithmetic was clearly more difficult than merely reporting each number. Overall it seemed to me more difficult to combine and reduce information than to transmit it. I designed a series of simple tasks that showed, under the conditions I had created, a certain amount of truth to the hypothesis that information reduction was a measure of thought, but the generality was too limited to be really useful and part way through my dissertation I realized that the goal of finding a new law was not to be realized.

Nonetheless, I persisted and was very delighted that my committee members, whom I greatly admired, were willing to let me pursue things. They must have known it would not succeed, but they never really conveyed their doubts. After working with numerous students myself, I know that sometimes it does pay to withhold criticisms and see what develops.

In an effort to extend my measurement of information reduction to the study of perception, I created stimuli of nine dots and applied various rules that distorted the dots and produced new patterns that were either close to

the original or very dissimilar. The distortions could be summarized by the amount of information change needed to produce them. Using the then popular methods of standard psychophysics, I could show that my information measure of distortions fit well with subjective ratings by people. Not surprisingly, when two patterns were heavily distorted it was very hard to classify them as members of the same category than if they were similar. That this finding could be summarized in terms of information reduction allowed my experiments to be seen as related by a single hypothesis.

The most enduring finding from the dot pattern work was based on studies that I conducted with Steve Keele after coming to Oregon. In my thesis I had taught people to call two very different patterns by the same name. We could easily extend this to study a type of concept or pattern learning. At the time of this research, studies of concept formation were dominated within psychology by an extremely rational approach based on ideas developed earlier by Aristotle and Mill. For example, Jerome Bruner had shown that Harvard undergraduates could reason about series of nonsense patterns consisting of conjunctions or disjunctions of distinct attributes. With appropriate feedback concerning whether or not the pattern reflected the concept, they were able to abstract an understanding of the rule that had been chosen. Clearly this work showed some of the reasoning capabilities of selected undergraduates, but concepts could be learned nearly automatically by young children, animals, and undergraduates, probably less reflective and more poorly trained than the Harvard students that Bruner had studied.

With only a vague understanding of the questions we were asking, Keele and I designed studies in which subjects learned to give a single name to four patterns highly distorted from a single prototype pattern in the manner used in my dissertation. When people memorized four such patterns, they showed the behavior of having a concept in that they gave a common response to what appeared to be very different looking events. However, we were pleased that they could do more than this. When we showed them the prototype which they had never before seen, they named it without difficulty and in recognition memory experiments they often said that they had actually seen it before.

I was very excited by these results. It seemed at the time that I had a new experimental demonstration not only of learning of schemas but of an automatic way of abstracting the essence from clues that would be as remote as would actually be found in the real world. This after all is what philosophers Berkeley and Mill had argued were abstract ideas derived from sensory input but capturing only what was most essential. Critics could and did argue that this was merely a kind of complex generalization process in which the prototype was not stored but simply recognized from storage of all of the exemplars. Indeed this criticism was made to my submission of a somewhat overheated paper, which we titled "On the Genesis of Abstract Ideas."

To my delight, and despite many criticisms, the paper was accepted by the editor to the *Journal of Experimental Psychology*, David Grant, who was thought to be as tough a critic as they come. He even accepted the title, which most of the referees wanted me to cut, with the comment this shows "we can trip the light fantastic."

This is not the end of the story. Thanks to the brilliance of Eleanor Rosch, my efforts ended up as a small part of what was certainly one of the intellectual triumphs of psychology during the late 20th century. Rosch first attempted to understand how color names were derived. The importance of this question derived from the general issue of whether names were really quite arbitrary or whether they depended critically upon the perceptual process. Benjamin Lee Whorf had argued that our perception or worldview was critically dependent upon the language we spoke; and how we perceived color had been made by Brown and others a central topic in testing whether perception depended on naming or not. Rosch was a student of Brown, and she conducted anthropological work with the Dani people, who apparently had only two color names (light and dark). She showed that they were better at learning to name colors that were excellent examples of our English color names (e.g., a clear, highly saturated red) than those close to the boundary. These results supported the idea that color names were not arbitrary but arose out of perceptual experience. Later, two Berkeley anthropologists built on this finding to show simple rules out of which perception led to naming that seemed to be general across cultures. In the era of cognitive neuroscience the dominance of biology for many of our cultural derivations (e.g., language) is not a surprise, but it was the first of the many important observations that Rosch made. Rosch argued for prototypical colors as the basis of classification and cited my studies, although, in fact, her work in no way really depended upon them. Nonetheless, I was delighted to have a role.

Rosch was to make even more important findings for the field of psychology. After all, not all trees and still less all games look alike, yet we have no trouble in forming these concepts. Rosch distinguished between basic concepts in which the exemplars looked similar (e.g., trees) and superordinate concepts (e.g., games) in which they did not. She found that with basic concepts there was an automatic extraction of the concept, and in studies that were built upon matching experiments she proved that the time to classify clear instances of a concept was really based on their distance from the prototype. After priming with the name animal, it was easier to match two good instances of animals (e.g., cow) than two instances equally familiar but which were not as close to the prototype (fish). Moreover, Rosch was able to write clearly that these studies were really violations of the overly rational idea of concept that had come into psychology via Bruner. Indeed, later writings on the issue often argued that Rosch had overturned

ideas of concepts that had begun with Aristotle and had shown instead how concepts were fuzzy or graded ideas based on simple derivations from our perceptual experience.

I thought at the time, and I still do, that Rosch's work represents an enduring and fundamental contribution to psychology for which she is due every credit that the field has. I came to see them as related to an observation made by Herbert Simon based on his concept of bounded rationality. We can, of course, learn to reason, but the thought processes reveal the concrete nature of how we do reason, whether in images or in concepts. Cognitive neuroscience has tended to confirm these ideas by showing how intimately frontal systems, highly abstracted from sensory input, are related to posterior systems such as the fusiform face area which are close to the perceptual input. Indeed, when dog and bird experts who were able to differentiate many different species were tested, they tended to activate the fusiform face areas just as nonexperts would do for human faces.

The work of Rosch has formed an important link between general social science and models of brain function. In my view one of the most important developments in the study of the higher mental processes has been based upon the importance of concrete analogical thinking in human reasoning. In a brilliant series of empirical and philosophical works, the linguist George Lakoff and the philosopher Mark Johnson have built upon the thinking of Rosch to argue that metaphors play a central role in how we think about movement, love, and politics. They argue that much of our knowledge is based upon what they call embodied reasoning. Again a clear reference to the kind of bounded rationality and concrete thought that was implied by Simon. However, in their hands we can see how so much of our thinking about the real world is governed by common and idiosyncratic metaphors. How the knowledge of the Munich agreement, ceding parts of Czechoslovakia to Germany, influenced the response to Vietnam and how our thinking about how the body moves influences how we regard covert shifts of attention.

While Rosch clearly made a major contribution toward how to conceive of human thought processes, my original dot patterns were employed in studies of human memory. During the 1970s and 1980s there were continual attacks on the idea of prototypes being abstracted from diverse inputs. Many researchers developed models in which, by storing each individual exemplar, it was possible to simulate the data obtained in my original memory experiments by methods that relied on simple stimulus generalization just as the referees of the original paper had suggested. I began to get the feeling that I was like the punch drunk fighter who had to be beaten up by each new model before it could be proposed for an important championship fight. The most creative and sustained of these models (Minerva) was developed by my friend and colleague Doug Hintzman. Minerva proposed a

creative way to store exemplars and use them in creating an echo that provided a basis for recognition of new instances. Thus, our data could be simulated even though there was never any storage of prototypes. Minerva was a great step forward in memory models and was really not very inconsistent with our way of thinking about prototypes. Because Minerva stored the results of the echo, and even though any given calculation was based on exemplars, once the calculation was made the prototype became a new exemplar and was now stored. The Minerva form of memory storage was sufficient to allow prototypes to stand for concepts and to play the role in metaphorical thinking that has been discovered.

Probably our work on prototypes would have died off with no resolution of the disputes except that Knowlton and Squire showed that patients whose memory had been impaired by brain lesions were at a great disadvantage in remembering exemplars but dealt very well with the prototype. These studies suggested that extraction of the prototype might not involve the medial temporal brain area found important for explicit storage. One way of examining this issue is to compare conditions when people are asked to explicitly recall an item with situations in which they can make use of the material but do not explicitly have to remember it. This task lent itself to fMRI studies, and they showed that implicit use of the primed word seemed to involve a portion of the right posterior cortex. In order to determine whether this activation represented an early priming by the stored information, I conducted a high-density electroencephalography (EEG) study with Rajendra Badigaiyan. We found that right posterior electrodes consistent with the fMRI activation differed between primed and unprimed words in the implicit condition during the first 150 milliseconds after input. These data suggested that right posterior activation of information was contacted automatically and rapidly after the input cue. On the other hand, activations in the explicit condition were mostly in hippocampal and frontal areas.

Another feature of the brain circuits related to expertise (including faces, word forms, and artificial and natural categories) is that in addition to the posterior area of activation they also involve frontal areas. In the case of visual words, for example, frontal areas including the left ventral frontal area and the anterior cingulate are active within 150 milliseconds after input, almost as fast as some of the posterior areas. In general, the frontal and posterior areas work together over a long time interval to integrate diverse information related to the problem solution. In the case of generating the use of a noun, which takes about 1100 milliseconds, the frontal areas are in communication with posterior areas related to semantics at 450 milliseconds. In general, brain studies have argued that there is close communication between frontal, posterior, and subcortical areas in generating the solution to problems. These findings provide a more objective basis for the role that concrete codes stored in visual and auditory areas may play in cognition.

Neural Basis of Mental Operations

Starting in 1978, I began to use a cue in an otherwise empty visual field as a way of moving attention to a target. We monitored eye movements and since only one response was required there was no way to prepare the response differently depending upon the cue. The results seemed to me to be very spectacular. We found that a covert shift of attention induced by a cue, presented as little as 100 milliseconds before a target, could enhance the speed of responding to target onset. We had trapped a covert attention shift and observed its movement. In fact, in one study response times to probes at intermediate locations were enhanced at intermediate times as though attention actually moved through the space. Whether attention moves through the intermediate space is still a disputed matter, suggesting the limitation of a purely behavioral study. At the time, it was also hard to conceive how a movement of attention could possibly be executed by neurons. Subsequently, Georgopoulos and colleagues showed that the population vector of a set of neurons in the motor system of a monkey could carry out what would appear, behaviorally, as a mental rotation. After that finding, a covert shift of attention did not seem too far-fetched.

About this time I became aware of a number of papers by Vernon Mountcastle and by Bob Wurtz using cellular recording to study the properties of cells in the posterior parietal lobe of the monkey. These papers suggested the possibility of attention cells in the parietal cortex that might be critically involved in orienting attention toward visual events. A Tuesday night meeting of our research group was assigned to read these papers. I asked whether our reaction time measures were the results of such attention cells. I thought, if the covert shifts of attention in humans could be connected with the monkey work, it might contribute to linking cognitive psychology to brain mechanisms. I don't think there was much enthusiasm for this idea at the time. After all, cognition was about software and what did it have to do with the parts of the brain in which cells were found in the monkey?

In 1979 I met Oscar Marin, an outstanding behavioral neurologist. He was about to move to Portland, Oregon, to set up a clinical and research effort at Good Samaritan Hospital, and he invited me to set up a neuropsychology laboratory in conjunction with the hospital. It was a perfect time for me because I had spent 6 months of 1979 in New York working with Michael Gazzaniga, whose career in psychology is probably familiar to most readers, and my brother Jerry, who helped me test patients with parietal lesions. I pursued these questions in the new laboratory in Portland. In the end I commuted to Portland once a week for 7 years. It was such a pleasure to work with Dr. Marin that the long drive was worthwhile.

The results seemed to me to be a revelation. Patients with different lesion locations in the parietal lobe, the pulvinar, and the colliculus all tended to show neglect of the side of space opposite the lesion, but in a

detailed cognitive analysis it was clear that they differed in showing deficits in specific mental operations of disengaging, moving, and engaging attention. As I saw it at the time, we had found a new form of brain localization. Different brain areas executed individual mental operations or computations such as disengaging from the current focus of attention (parietal lobe), moving or changing the index of attention (colliculus), and engaging the subsequent target (pulvinar). No wonder Lashley thought the whole brain was involved in mental tasks. It was not the whole brain, but a widely dispersed network of quite localized neural areas. Even looking back from the perspective of 20 years, I can again feel the excitement I had surrounding this idea at the time.

I read an article in *Scientific American* indicating changes in cerebral blood flow in the brain when reading silently. In cognitive psychology, reading had been studied quite a lot and we knew something about the orthographic, phonological, and semantic operations that must have taken place while reading, but they would be combined in the overall blood flow. Even more compelling for the possible anatomy of mental operations was a paper appearing in 1985 by Per Roland indicating that different parts of the brain were active during way finding, mental arithmetic, and verbal tasks. However, even in this paper there was no effort to uncover the specific operations that might be performed by the brain areas involved.

About this time Washington University School of Medicine started a national search for a psychologist who might work in conjunction with the developing PET center led by Marc Raichle. It might be surprising to people how reluctant psychologists were to take a chance on brain imaging. For me this was the opportunity to test the idea that arose from the neurological studies that individual mental operations would be localized in separate brain tissue. James S. McDonnell had wanted to develop an Institute that would study extrasensory perception, but the powers that be at Washington University were not going to do that. Instead, they agreed to a Center for Higher Brain Function. A psychologist who studied brain function was about as mystical as they wanted to go, and Marc Raichle and his colleagues at Washington University recognized the importance of being able to use PET to illuminate questions of higher mental function.

I had gone to St. Louis in the hopes of pursuing work on attention. When I talked to neurologists about covert shifts of attention (without eye movements) and then proposed to break the invisible shift into component operations like disengaging and moving, I saw eyes glaze and interest wane. Language studies have the advantage that the operations were more concrete and that neurosurgeons valued knowledge about the localization of language areas to aid them in avoiding such areas during surgery. Our language studies, summarized in my book with Marc Raichle, had an important influence on the field, fostering many studies of language and other cognitive processes. The development of fMRI allowed studies that were less

invasive and more precise, and they have generally confirmed and greatly extended our language studies and shown that in many other areas of cognition a network of widely scattered brain areas become active, which when orchestrated together allow us to carry out a variety of tasks.

The St. Louis studies did quite a lot for the development of neuroimaging and, in the main, supported the idea that widely scattered brain areas were involved when any task was studied. Some people have thought that these areas were specific for domains of function like language or face stimuli, and so on. I have maintained the importance of mental operations, without denying that domain specificity may also play a role in understanding localization. In the area of face processing, for example, there has been a lot of dispute over whether there is a specific face area because experts in other domains activate the same area when thinking about their domain of expertise. However, if one thinks about localization of mental operations, it seems clear that faces and other objects, where we come to recognize the individual people or objections via their distinctive features, share operations in common. A similar argument can be applied to the visual word form area.

Attention is one of the areas that has been widely studied by fMRI. Maurizio Corbetta came to St. Louis after I had left and he, together with Gordon Shulman, have provided strong evidence for localization of quite separate mental operations within two areas of the parietal lobe that form a portion of a larger network whose functions are to align attention with the target. Although my initial speculation of which operations occurred in which areas was not entirely correct, the beautiful localized brain areas support the overall localization hypothesis.

Development of Attention

When I returned to Eugene from St. Louis in 1989, it was with the goal of forwarding research in two directions. It was my conviction that the findings we made in St. Louis on processing visual and auditory words would lead to evidence for many networks of brain areas involved in mental activity. While I did not dream that the work would grow to the magnitude that actually took place, following the use of fMRI, the general shape of what has happened seemed clear right from the start. I also realized that the skills needed to improve localization and understand the anatomy of cognition were far removed from what I could do best. Instead I concentrated on measuring the operation of these networks in real time.

Even before leaving St. Louis, I teamed with Avi Snyder and Marc Raichle to carry out studies of the time course of processing visual words and nonsense material using the 16-channel EEG system available there. However, in Eugene, Don Tucker was developing a new way of taking EEG from the skull, which would allow many electrodes to be put into place at once. Together we set up a laboratory in Straub Hall and began to compare

his 32-channel EGI system with a 32-channel electro-cap using the Grass amplifiers from older studies. The results convinced me that there was no great loss from the relatively high impedances used in the EGI system. We were able to see scalp signatures of brain areas corresponding to the major generators found in the previous PET and later fMRI language studies. In my work with Yalchin Abdullaev and Bruce McCandliss, among others, it was exciting to see activity in the visual word form area, anterior cingulate and frontal and temporo-parietal cortex in real time, as the network computed different aspects of the language tasks we used. More recently, real-time fMRI allows some of this work to be done using that method, but because of the speed of mental processes and the importance of brain rhythms in connecting neural areas, EEG remains an important contributing method.

The study of attention has been a central topic from the start of human experimental psychology. However, many who write about attention seem to view it as a slightly mystical issue not amenable to scientific definition. They prefer to think of it as an emerging property, usually of sensory or motor systems and not an organ system with its own anatomy, evolutionary history, and function. In 1972, I began work on attention by dividing the overall system into components. It was not until 1990 when, due largely to the development of neuroimaging in humans and cellular recording in nonhuman primates, it became possible to outline the neural networks that underlie these functions.

In our work, we laid out networks of brain areas related to obtaining the alert state, orienting to sensory stimuli, and executive control involve in resolving conflict between other brain networks. I have written extensively about the anatomy, circuitry, neuromodulators, and genes related to common properties and individual differences in network efficiency. As a result I have come to regard attention as an organ system consisting of separate networks carrying out its various functions. Attention networks are of special importance because attention is involved in many functions, including our ability to control our behavior in the face of conflicting external and internal demands. During the late 1990s, as the result of a donation by the Mortimer Sackler family, I had the unique opportunity to set up an Institute for Developmental Psychobiology in New York City. My application was aided by the fact that my brother headed the Neuro-oncology Department right across the street. Moreover, there was a clear need for students trained to understand the developing human brain, and the presence of scanning and other facilities in New York was attractive. During my stay in New York, it was possible to recruit B. J. Casey and Bruce McCandliss to the Sackler Institute faculty, and they have given it a leadership role in pediatric neuroimaging and in the new area of brain and education.

According to theories developed by Mary Rothbart, a control system develops in late infancy or early childhood that involves attention and can

also be measured through parental reports as a higher-order variable she called effortful control. Since the executive attention system involves the anterior cingulate, which is well activated by conflict, we developed a series of conflict tasks that could be executed by children. Scanning with fMRI in New York, Jin Fan was able to show that these various conflict tasks activated a common network, including the anterior cingulate gyrus.

I never expected to study individual differences. Cognitive psychology was about mental processes common among humans, and I thought one of the most important results of the early neuroimaging studies was evidence that the data of subjects could be averaged to provide a common anatomy even for higher mental tasks where one might have suspected a lot of individuality. Of course the early PET data had a large area of blur in activity due to averaging, so, despite evidence for commonality, there was also opportunity for individuality. The development of fMRI has allowed enough data to be collected on an individual that activity can be plotted in relation to the anatomy of the individual brain, thus opening the way for more detailed analysis of individuality in brain activation.

Two things tipped me in the direction of wanting to study individual differences. First was the influence of Mary Rothbart's elegant theory of the role of effortful control in child socialization. Our data showed that all during childhood there were correlations between the executive attention network measured in cognitive tasks and parent-reported effortful control. This seemed remarkable evidence that individual differences in laboratory experiments were important enough to relate to the diverse behaviors of everyday life that would be obvious to parents.

The second reason was due to meeting John Fossella and Tobias Sommer at the Sackler Institute in New York. They both had been trained in molecular biology, John to the Ph.D. level, and were eager to explain to me some of the new opportunities opened up by the human genome program. With the support of Mary and her team in Eugene, we set out to measure adult individual differences both by questionnaire and experimental test and to determine how they might be influenced by genetics. John Fossella took cheek swabs from 200 New York adults and using the magic of polymerase chain reaction (PCR) was able to genotype them for dopamine genes. We found that individual differences in the attentional network that related to effortful control were influenced by dopamine genes, and an fMRI study, conducted by Jin Fan, B.D. showed that the genetic differences influenced the degree of brain activation in the anterior cingulate.

Sometimes the effort to relate genes to brain networks and complex behavior is taken as evidence that everything is hard wired and that experiences are unimportant. This of course is not at all true. In a study of 2-year-old children we found, for example, that children with the 7-repeat allele of the DRD4 gene were greatly influenced in critical aspects of behavior by the quality of parenting. However, for children without the 7-repeat allele,

parenting made little difference in these same aspects of their behavior. The results with the DRD4 gene were particularly important, because they led us to the hypothesis that genetic variation might be selected in human evolution because they made the child more influenced by culture factors like parenting.

Mary Rothbart attended a meeting at NIH and reported that Duane Rumbaugh and David Washburn had trained monkeys in a version of the Stroop effect. Since conflict tasks such as the Stroop served as a marker of executive attention, we decided to adopt their methods to train preschool children. With Charo Rueda, we developed a training program based on the animal research and randomly assigned 4- and 6-year-old preschool children for 5 days of training in this program or to a control group using interactive videos. We found clear improvement in the executive network at both ages. Subsequently, Rueda showed that this small amount of training still had an effect after 2 months. At the time we did these experiments it seemed unlikely that such a small amount of attention training could improve brain networks, but subsequently there has been quite a lot of research showing improvement in attention with different methods. In 2007 Mary Rothbart and I published our volume *Educating the Human Brain*. It reviews our efforts to train attention in preschool children, and we still hope it may help lead to improved experience for children in preschool that might influence their later education.

Most of the work on training has involved practice with a specific brain network. In our case we tried to systematically exercise the brain's executive attention network. However, more recently I have joined with Dr. Yiyuan Tang in proposing a second method which might be useful. In this method we attempt to train the brain state. Brain states are being studied a lot these days, thanks to Marc Raichle showing the importance of the default state. In our studies, we used a version of meditation adapted from traditional Chinese medicine (integrated body-mind training [IBMT]). With U.S. and Chinese undergraduates randomized to IBMT or to a relaxation training control condition, improvements were found in executive attention and in cortisol secretion to a cognitive challenge. These changes appeared to involve a combination of central and autonomic (parasympathetic) system enhancements. In the case of IBMT there are no specific exercises involving practice of attention. Instead one attempts to induce a highly relaxed but mentally focused state.

Currently we continue training studies and are examining how attention networks are shaped by genes and experience in children from 7 months to 4 years of age. Examining individual differences may seem a somewhat surprising method to identify genes involved in the development of common brain networks. However, I believe that those genes related to individual differences in the efficiency of a particular network will prove to be the same as those involved in building the general network. Future research will determine how sound this logic is.

It is now over 50 years since I first began to study psychology. At the time I began my work in my fondest dreams I could not have imagined that we would be discussing the mechanisms of voluntary control, and using neuroimaging to determine which brain areas are involved and genetic variation to examine how those brain areas were developed. When I married in 1958 and my wife asked why I wanted to go to graduate school, I told her, "To figure out how the brain works." Although we have a long way to go, it seems to me that in these 50 years I had a wonderful opportunity to help make a start.

Selected Bibliography

- Abdullaev, Y.G., & Posner, M.I. (1998). Event-related brain potential imaging of semantic encoding during processing single words. *NeuroImage*, 7: 1–13.
- Badgaiyan, R., & Posner, M.I. (1998). Mapping the cingulate cortex in response selection and monitoring. *NeuroImage*, 7: 255–260.
- Fan, J., Fossella, J.A., Summer T. Wu, Y., & Posner, M.I. (2003) Mapping the genetic variation of executive attention onto brain activity. *Proceedings of the National Academy of Science USA* 100: 7406–7411.
- Fan, J., McCandliss, B.D., Sommer, T., Raz, M., & Posner, M.I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 3(14): 340–347.
- Posner, M.I. (1978). *Chronometric Explorations of Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Posner, M.I. (1980). Orienting of attention. The 7th Sir F.C. Bartlett Lecture. *Quarterly Journal of Experimental Psychology*. 32: 3–25.
- Posner, M.I., & Boies, S.J. (1972). Components of attention. *Psychological Review*, 78: 391–408.
- Posner, M.I., & Keele, S.W (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77: 353–363.
- Posner, M.I., & Mitchell, R.F. (1967). Chronometric analysis of classification. *Psychological Review*, 74: 392–409.
- Posner, M.I., & Petersen, S.E. (1990) The attention system of the human brain. *Annual Review of Neuroscience*, 13: 25–42.
- Posner, M.I., & Raichle, M.E. (1994). *Images of Mind*. New York: Scientific American Books.
- Posner, M.I., & Rothbart, M.K. (1998). Attention, self regulation and consciousness. *Philosophical Transactions of the Royal Society of London B*, 353: 1915–1927.
- Posner, M.I., & Rothbart, M.K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, 12: 427–441.
- Posner, M.I., & Rothbart, M.K.(2007). *Educating the human brain*. Washington DC:APA Books.
- Posner, M.I., Boies, S.W., Eichelman, W., & Taylor, R. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monography*, 79: 1–16.

- Posner, M.I., Petersen, S.E., Fox, P.T., & Raichle, M.E. (1988). Localization of cognitive functions in the human brain. *Science*, 240: 1627–1631.
- Posner, M.I., Rothbart, M.K., & Sheese, B.E. (2007). Attention genes. *Developmental Science*, 10: 24–29.
- Posner, M.I., Walker, J.A., Friedrich, F. J., & Rafal, R.D. (1984). Effects of parietal lobe injury on covert orienting of visual attention. *Journal of Neuroscience*, 4: 1863–1874.
- Raz, A., Fan, J., & Posner, M.I. (2005) Hypnotic suggestion reduces conflict in the human brain. *Proceedings of the National Academy of Sciences USA*, 102: 9978–9983.
- Rueda, M.R., Rothbart, M.K., McCandliss, B.D. Saccamanno, L., & Posner, M.I. (2005). Training, maturation and genetic influences on the development of executive attention *Proceedings of the National Academy of Sciences*, 102: 14931–14936.
- Sheese, B.E., Voelker, P.M., Rothbart, M.K., & Posner, M.I. (2007). Parenting quality interacts with genetic variation in Dopamine Receptor DRD4 to influence temperament in early childhood *Developmental and Psychopathology*, 19: 1039–1046.
- Tang, Y.Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., Yu, Q Sui, D., Rothbart, M.K., Fan, M., & Posner, M.I. (2007). Short term meditation training improves attention and self Regulation. *Proceedings of the National Academy of Sciences USA*, 104: 17152–17156.
- Voelker, P. Sheese, B.E., Rothbart, M.K & Posner, M.I., & Rothbart, M.K. (2009). Variations in COMT gene interact with parenting to influence attention in early development *Neuroscience*, 164(1): 121–130.