Johnson-Laird, P. N. (1998). Imagery, visualization, and thinking. In J. Hochberg (Ed.), Perception and Cognition at the Century's End (pp. 441-467). San Diego, CA: Academic Press.

CHAPTER 14

Imagery, Visualization, and Thinking

P. N. Johnson-Laird

Alt our ideas and concepts are only internal pictures.
—Ludwig Boltzmann (1899)

I. INTRODUCTION

What is the relation between the ability to see the world and to think? For the Renaissance humanist, Leonardo da Vinci, the two processes were almost one and the same: to visualize *was* to think. The eye was the instrument of thought, and the artist's ability to make pictures provided a special medium in which to carry out "thought experiments." This view, however, has been supplanted by the development of quantum mechanics and other abstract disciplines in which the objects of thought are all but impossible to visualize. The manipulation of abstract representations—if it occurs—does indeed seem to be a mode of thinking that is not visual. The emphasis on abstraction and mathematics can be found in many intellectual disciplines—from logic to mechanical engineering—in which practitioners concomitantly play down the role of visualization. These pedagogical trends, however, may have little basis in fact. The aim of the present chapter is accordingly to reconsider the nature of mental representations and thinking, that is, of how the mind's eye may help the mind's mind. It will use the results of psychological experiments to reach some new conclusions about the relations between visualizing and thinking.

Section II of the chapter begins with the traditional account of visual imagery—the introspective reports of individuals claiming to use images, and the commentators' claims—according to which visualization is central to the creation of novel ideas. One source of renewed interest in the topic was the rediscovery of imagery by psychologists. After years of deliberate neglect during the Behavioristic era, cognitive

Perception and Cognition at Century's End

Friedrich August von Kekulé, for example, described how in 1865 the ring-like structure of benzene came to him in a dream:

> I turned my chair to the fire and dozed. Again the atoms were gamboling before my eyes. This time the smaller groups kept modestly in the background. My mental eye, rendered more acute by repeated visions of this kind, could now distinguish larger structures, of manifold conformation; long rows, sometimes more closely fitted together; all twining and twisting in snakelike motion. But look! What was that? One of the snakes had seized hold of its own tail, and the form whirled mockingly before my eyes. As if by a flash of lightning I awoke. (Findlay, 1937, p. 43)

The snake biting its tail had given him the clue to the puzzle.

Other scientists are said to have thought in terms of images. Gruber (1974, p. 237) reports that Darwin's notebooks are full of images, though Darwin himself wrote that he found it just as easy to think about abstract ideas as concrete ones. It is worth noting that Darwin's "thought experiments" in *The Origin of Species* do not hinge specifically on visualizable events, for example:

> It is good thus to try in our imagination to give any form [of plant or animal] some advantage over another. Probably in no single instance should we know what to do, so as to succeed. It will convince us of our ignorance on the mutual relations of all organic beings; a conviction as necessary, as it seems to be difficult to acquire. All that we can do, is to keep steadily in mind that each organic being is striving to increase at a geometrical ratio; that each at some period of its life, during some season of the year, during each generation or at intervals, has to struggle for life, and to suffer great destruction. (Darwin, 1859/1968, p. 129)

Darwin was a subject of Galton's well-known questionnaire study of imagery among Fellows of the Royal Society (Galton, 1880/1928), which revealed—somewhat to Galton's consternation—that many of them claimed to think without using images. Yet most scientists who have written on their own thought processes have emphasized the role of imagery. Perhaps the best known testimonial occurs in Einstein's letter to Hadamard (1996):

> The words of the language, as they are written or spoken, do not seem to play any role in my mechanism of thought. The psychical 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.
>
> There is, of course, a certain connection between those elements and relevant logical concepts. It is also clear that the desire to arrive finally at logically connected concepts is the emotional basis of this rather vague play with the above mentioned elements. But taken from a psychological viewpoint, this combinatory play seems to be the essential feature in productive thought—before there is any connection with logical construction in words or other kinds of signs which can be communicated to others. (p. 142)

Max Wertheimer, one of the founders of Gestalt psychology, had occasion to interview Einstein, and corroborated the role of visualization (Wertheimer, 1961, p. 228;

psychologists took up the topic again. The chapter describes some of their key findings. No sooner had images been rediscovered than certain skeptics, notably the Canadian cognitive scientist, Zenon Pylyshyn, argued that they are epiphenomenal, that is, they play no causal role in mental life. If these skeptics are right, the emphasis on visualization in thinking is misplaced. The real work is done by so-called "propositional representations," that is, representations of propositions in a mental language. Hence, we need to establish whether images, propositional representations, or some other sort of mental representations underlie thought.

With this aim, Section III of the chapter turns to a test case: reasoning. Its mechanisms are not accessible to introspection, but they can be characterized theoretically. The chapter describes the orthodox theory that reasoning depends on propositional representations and formal rules of inference like those of a logical calculus—a view advanced by most cognitive scientists. It then outlines an alternative theory that reasoning is not a formal process at all, but a semantic process that depends on the manipulation of "mental models" of states of affairs according to the fundamental principle of validity, an argument is valid if its conclusion must be true given that its premises are true. The theory yields a surprising prediction: certain inferences should be like illusions (i.e., they will have conclusions that arc compelling, yet that are completely wrong). A recent discovery is that these illusory inferences do exist.

What is the relation between mental models and mental images? Section IV of the chapter argues that models underlie the experience of imagery, but models themselves may contain elements that cannot be visualized, such as an annotation representing negation. Experiments confirm the existence of such annotations, and they show that reasoning is unaffected by the "imageability" of the materials. The operations that can be directly carried out on images correspond to visual transformations rather than to deep conceptual processes. Even mental rotations of images representing objects are likely to depend on an underlying three-dimensional model. Images, however, may have a symbolic function, and thus diagrams can help people to reason about entities that cannot be readily visualized.

Finally, the chapter draws some morals about visualization (in Section V). The result is a rehabilitation of imagery in the face of the skeptics, but a limitation on imagery in the face of its more ardent adherents.

II. VISUALIZATION

A. Introspections about Visualization

The aim of this section is to review visual imagery. According to many commentators, it is fundamental to scientific and technological invention (e.g., Ferguson, 1977; Miller, 1984; Shepard, 1988; Valéry, 1894). The claim is based in part on the traditional views of scientists themselves, particularly 19th century physicists (see, e.g., Boltzmann, 1899). It is also bolstered by a number of celebrated anecdotes.

for corrections to other aspects of his account, which are historically inaccurate, see e.g., Miller, 1984, Ch. 5). Gleick (1992, p. 131) has similarly emphasized the role of imagery in the late Richard Feynman's thinking. And even one psychologist, Feldman (1988), claims that an image played a significant part in his own thinking.

B. The Rediscovery of Images

Mentalistic psychologists studied visual imagery in the last century and the early years of this century (e.g., Binet, 1894; Perky, 1910). But the topic fell into disrepute as a result of a dispute, the so-called "imageless thought" controversy. A group of psychologists at Würzburg led by Karl Marbe and Oswald Külpe claimed that their subjects often reported a kind of conscious but unanalyzable experience that was neither an image nor an awareness of an intention or act of will. These *Beswusstseinslagen,* or "imageless thoughts," ran contrary to the Aristotelian view of thinking as an association between ideas. They were taken to imply the existence of unconscious processes leading to their appearance in consciousness. To base a theoretical argument solely on introspections, however, was an egregious error. It is impossible to establish their authenticity—even Kekulé's dream, for example, may turn out to have been a fraud (see Wotiz & Rudofsky, 1984, cited by Gruber, 1994). Wundt, the leading psychologist of the day, challenged the Wiirzburg school, as did his student, Titchener, one of the founders of American psychology, who declared that his thinking was always accompanied by imagery (see Humphrey, 1951, for an authoritative history). The controversy was never settled; rival theorists traded rival introspections, but it was swept away, along with the study of imagery, by the rise of Behaviorism.

With the revival of mentalism and the cognitive revolution in psychology, psychologists rediscovered imagery (Holt, 1964). They were soon to distinguish between what they termed verbal and visual representations (Bower, 1970; Paivio, 1971). With Shepard's studies of the mental rotation of visual images (e.g., Shepard & Metzler, 1971), the topic appeared to have been rehabilitated within psychology. Shepard and his colleagues demonstrated that individuals can transform objects mentally in a variety of ways. In the first of their experiments, which itself was suggested by a spontaneous kinetic image in one of Shepard's dreams, the subjects saw two drawings of a "nonsense" figure assembled out of ten blocks glued together to form a rigid object with right-angled joints. Their task was to decide whether the pictures depicted one and the same object from different points of view. The time they took to make their decision increased linearly with the angular difference between the orientations of the two objects. This result held both for rotations in the picture plane and for rotations in depth. It implied that subjects could mentally rotate their representation of such objects at a rate of about 60° per second.

Kosslyn and his colleagues obtained similar results when they asked their subjects to scan from one landmark to another in their image of a map that they had committed to memory (see, e.g., Kosslyn, Ball, & Reiser, 1978). Kosslyn (1980) also estimated the size of the mental "screen" on which images are projected: subjects had to form an image of, say, an elephant and then imagine walking toward it until the image began to overflow their minds' eye. They stopped further from their image of an elephant than from their image of a smaller animal, such as a dog. The size of the mental screen is about the same for an image as it is for a visual perception. Many other investigations of imagery, from its mnemonic value (Luria, 1969) to its special storage in short-term memory (Baddeley & Hitch, 1974), seem to imply that visual images are a distinct medium of mental representation.

This view was challenged by Pylyshyn (1973). He argued that a distinct medium of representation would be part of the functional architecture of the mind and so its properties could not be affected by an individual's beliefs or attitudes. The case is comparable, he claimed, to the architecture of a digital computer: the design of its hardware cannot be modified by a program that the computer is running. Mental architecture is thus "cognitively impenetrable," whereas imagery is easily influenced by an individuals' beliefs. Indeed, Pylyshyn argued, the results of the rotation and scanning experiments might merely reflect the ability of subjects to simulate how long it would take to rotate an actual object, or to scan across an actual map. Such simulations would reveal nothing about the real nature of mental representations. In Pylyshyn's view, the mind depends on formal computations carried out on a single sort of representation: syntactically structured "propositional representations" expressed in a mental language. Images undoubtedly occur as subjective experiences, but they are epiphenomenal (i.e., they do not play any causal role in mental processes).

There are two ways to resolve the argument between imagists and propositionalists—the "thoughtless imagery" controversy, as it might be dubbed (Johnson-Laird, 1983, Ch. 7). In one sense of "propositional representation," the propositionalists must be right. All mental life depends on the brain's "machine code," and so everything must be reduced to nerve impulses and synaptic events, just as all computations no matter how complex can be reduced to the shifting of bits from one computer register to another. In another sense of "propositional representation," the imagists may be right, and there is a real distinction between images and propositional representations. In this sense, they are both high-level representations within the same computational medium, just as lists of symbols and arrays of symbols are distinct representations within a high-level programming language, such as LISP. If we draw the distinction in this way, we are left with an empirical question: What sorts of high-level mental representation does human thinking depend on? The aim of Section III is to answer this question by considering the psychology of reasoning.

III. THE PSYCHOLOGY OF REASONING

A. The Theory of Formal Rules and Propositional Representations

Consider the following problem about a particular hand of cards:

1. There is a king in the hand, or there is an ace in the hand, or both.
2. There is not a king in the hand.
3. What follows?

mal rule theories, and most investigators took *for* granted the existence of a mental logic. Yet an alternative did exist: reasoning could be based on a semantic method rather than the syntactic method of formal rules. In logic, a comparable distinction is drawn between "proof-theoretic" methods based on formal rules and "model-theoretic" methods based on semantic principles. The next section outlines a psychological theory based on a semantic method.

B. The Theory of Mental Models

The physicist Ludwig Boltzmann (1890) wrote of scientific thinking in the following terms:

> The task of theory consists in constructing an image of the external world that exists purely internally and must be our guiding star in thought and experiment; that is in completing, as it were, the thinking process and carrying out globally what on a small scale occurs within us whenever we form an idea. (p. 33)

The Scottish psychologist and physiologist, Kenneth Craik, similarly conceived of thinking in terms of the following programmatic idea:

> It the organism carries a "small-scale model" of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and the future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it. (Craik, 1943, Ch. 5)

Mental models can be constructed on the basis of visual perception (Marr, 1982) or verbal comprehension (Johnson-Laird, 1983). Their essential characteristic is that their structure corresponds to the structure of what they represent. Like a diagram (Maxwell, 1911) or an architect's model, the parts of the model correspond to the relevant parts of what it represents, and the structural relations between the parts of the model are analogous to the structural relations in the world. Hence, a model represents a set of individuals by a set of mental tokens, it represents the properties of the individuals by the properties of the tokens, and it represents the relations among the individuals by the relations among the tokens. And like a diagram, the model is partial because it represents only certain aspects of the situation. There is therefore a many-to-one mapping from possible states of affairs to the model. Images have these properties, too, but as we shall see, models and images differ from one another. Models need not be visualizable and, unlike images, they may represent several distinct sets of possibilities.

The theory of mental models postulates that human reasoners who have no logical training represent states of affairs using mental models. Psychologists cannot directly inspect mental models, and so the evidence for their existence and format is indirect. In the case of reasoning—deductive and inductive—the model theory

Most people respond rapidly with the correct conclusion:

There is an ace in the hand.

How did you carry out this inference? Introspection alone cannot tell you. Psychologists have studied reasoning since the turn of the century, and Störring (1908) found that his subjects reported using either images or verbal methods to reason. But the dominant view these days is that the mind is equipped with formal rules of inference, which it uses in order to reason, either deductively or inductively—a view that implies that visualization plays no role in reasoning.

In general, *formal rule* theories, as I shall henceforth refer to them, postulate that reasoners construct propositional representations of premises, identify their logical structure, apply formal rules of inference one at a time in a chain of steps that leads from the premises to the conclusion, and express this conclusion with its appropriate linguistic content (see, e.g., Draine, Reiser, & Rumain, 1984, Rips, 1994). Thus, the logical structure of the example above matches the formal rule of inference:

p or q, or both
not-p
∴ q,

which yields the conclusion. Formal rule theories postulate separate rules of inference for each of the main logical connectives: "if", "and", and "or". Table 1 summarizes the main rules. The theories predict that the greater the number of steps in a derivation, the harder the deduction should be, but they allow that certain rules may be harder to use than others.

One disquieting phenomenon for formal rule theories is that the content of the premises can have a striking effect on deductive performance (Wason & Johnson-Laird, 1972). Twenty years ago, however, there appeared to be no alternative to for-

TABLE 1 Some Typical Formal Rules of Inference Postulated as Part of Mental Logic by Many Psychologists

Rules that eliminate connectives	Rules that introduce connectives
p & q ∴ p	p q ∴ p & q
p or q not-p ∴ q	p ∴ p or q
if p then q p ∴ q	p ⊢ q ∴ if p then q (where "⊢" signifies that q can be derived from hypothesizing p)

postulates that reasoners construct a model, or set of models, based on the meaning of premises, perception, and any relevant general knowledge. They formulate a conclusion by describing a relation in the models that was not explicitly asserted by any single premise. Finally, they attempt to assess the strength of the inference. Its strength depends on the believability of its premises and on the proportion of models of the premises in which the conclusion is true (Johnson-Laird, 1994). The theory accordingly provides a single psychological mechanism for reasoning about necessary, probable, and possible conclusions:

1. A conclusion that holds in all possible models of the premises is *necessary* given the premises (i.e., it is deductively valid).
2. A conclusion that holds in most of the models of the premises is *probable.*
3. A conclusion that holds in at least one model of the premises is *possible.*

To illustrate reasoning by model, reconsider the earlier example:

1. There is a king in the hand, or there is an ace in the hand, or both.
2. There is not a king in the hand.
3. What follows?

The first premise calls for a set of models that represent the three possibilities, shown here on separate lines:

a
k a

where "k" denotes a king in the hand, and "a" denotes an ace in the hand.

A crucial assumption of the model theory is that individuals normally minimize the load on working memory by representing explicitly only those contingencies that are true (Johnson-Laird & Byrne, 1991). The models above, for example, do not make explicit what is false (i.e., an ace does *not* occur in the first model and a king does *not* occur in the second model). Reasoners should make a mental "footnote" to this effect, which can be used to make the models wholly explicit if necessary, but the theory assumes that footnotes are rapidly forgotten. For many deductions in daily life, there is no need to make models completely explicit. The premise,

The king is not in the hand,

rules out two of the models in the initial set, and all that is left is a single model:

which supports the conclusion,

There is an ace in the hand.

No other model of the premises refutes this conclusion, and so it is valid.

The theory postulates that a conditional, such as:

If there is a king in the hand then there is an ace in the hand is initially represented by the following two models:

In this case, people realize both cards may be in the hand, which they represent in an explicit model, but they defer a detailed representation of the case where the antecedent is false (i.e., where there is *not* a king in the hand), which they represent in a wholly implicit model denoted here by an ellipsis. Reasoners need to make a mental footnotes that a king cannot occur in the hands represented by the ellipsis, whereas an ace may, or may not, occur in these hands. Once again, the theory assumes that footnote are rapidly forgotten. Table 2 summarizes the models for the major connectives, and it also shows the fully explicit models, which do represent the false contingencies. There are occasions where reasoners do represent them.

The theory may seem to be no more than a notational variant of formal rule theories, but in fact it makes predictions that cannot be made by them. First, given the mind's limited processing capacity, the theory predicts that the greater the number of models that have to be constructed to make a deduction, the harder the task should be. More models mean more work. Second, the theory predicts that erroneous conclusions should occur because reasoners sometimes overlook possible models of the premises. Third, the theory predicts that certain inferences will be illusory (i.e., they will have conclusions that are compelling but completely wrong). In the following sections, we will review the experimental evidence for these predictions.

TABLE 2 The Sets of Models for the Main Sentential Connectives[a]

Connective	Initial Models		Fully Explicit Models	
A and B:	A	B	A	B
A or else B:	A		A	¬B
		B	¬A	B
A or B, or both:	A		A	¬B
		B	¬A	B
	A	B	A	B
If A then B:	A	B	A	B
			¬A	B
			¬A	¬B
If, and only if, A then B	A	B	A	B
			¬A	¬B

[a] The central column shows the initial models that the theory postulates for human reasoners, and the right-hand column shows the fully explicit sets of models. The symbol "¬" denotes negation, and the symbol " . . . " denotes a wholly implicit model. Each line represents a separate model.

The model theory predicts that the first problem, which calls for one model, should be easier than the second problem, which calls for at least two models. The process of constructing such models from verbal assertions has itself been modeled in a computer program (see Johnson-Laird & Byrne, 1991).

Formal rule theories of spatial reasoning, such as those proposed by Ohlsson (1984) and Hagert (1984), make exactly the opposite prediction. The first problem requires a formal derivation to establish the relation between the plate and the cup, which is then used to derive the relation between the fork and the knife. The second problem does not require a formal derivation of the relation between the plate and the cup, because the second premise directly asserts it. The second problem has a derivation that is just part of the derivation of the first problem, and so according to the formal rule theories it should be easier than the first problem.

Byrne and Johnson-Laird (1989) presented their subjects with sets of one-model and multiple-model spatial inferences, and a further set that did not support a valid answer. The subjects did all three sorts of problem, which were presented in a random order. They made 70% correct responses to the one-model problems, but only 46% correct responses to the multiple-model problems. Their correct conclusions were also reliably faster to the one-model problems (a mean of 3.1 sec) than to the multiple-model problems with valid answers (3.6 sec). In the example above, however, the multiple-model problem has an irrelevant first premise, and so a second experiment examined one-model problems with an irrelevant first premise. The results were the same: one-model problems were reliably easier than multiple-model problems.

One adherent of formal rules has argued that the trouble with these experiments is that the subjects were asked to imagine the objects on a table top, and that this instruction "obviously biased the subjects to use an imaginal strategy that favored the mental-model predictions (or placed a strong task demand on them to respond as if they were trying to image the arrays)" (see Rips, 1994, p. 415). This argument is unlikely to be correct, because the results have been replicated for problems based on temporal relations, such as,

a before b.
b before c.
d while a.
e while c.
What's the relation between d and e?,

where "a," "b," etc., stand for everyday events, such as "John shaves," "he drinks his coffee," and so on (see Schaeken, Johnson-Laird, & d'Ydewalle, 1996). In these studies, the subjects were obviously *not* told to imagine a table top; and, as in the spatial studies, they were given no instructions about how to do the task. The evidence accordingly corroborates the model theory and runs counter to the formal rule theories.

C. The Evidence for Mental Models

1. Mental Models and Spatial Reasoning

The ability to reason about spatial relations is likely to depend on the construction of mental models. Byrne and Johnson-Laird (1989) investigated the following sort of problems, which describe the layout of objects on a table top:

1. The cup is on the right of the spoon.
2. The plate is on the left of the spoon.
3. The knife is in front of the cup.
4. The fork is in front of the plate.
5. What's the relation between the fork and the knife?

The premises call for the model:

plate spoon cup
fork knife

where the left-to-right axis represents the left-to-right axis in the world, and the vertical axis represents the front-to-back axis in the world. Reasoners may well visualize the shapes of the various utensils. The model yields the answer to the question:

The fork is on the left of the knife.

No model of the premises refutes this conclusion, and so it is valid.

Now consider the following problem in which one word in the second premise has been changed:

1. The cup is on the right of the spoon.
2. The plate is on the left of the cup.
3. The knife is in front of the cup.
4. The fork is in front of the plate.
5. What's the relation between the fork and the knife?

The premises are spatially indeterminate because they are consistent with at least two distinct layouts:

plate spoon cup
fork knife

spoon plate cup
fork knife

In either case, however, the same conclusion follows as before:

The fork is on the left of the knife.

Inferences cannot be based on models alone. When the deductive procedure searches for an alternative model of the premises to refute a putative conclusion, it needs access to an independent representation of the premises. Consider, for example, the following model:

| ○ | △ | ⋆ |

and the putative conclusion:

The circle is on the left of the star.

The following model refutes the conclusion:

| * | A | O

but it will be relevant only if it is also a model of the previous premises, and the model itself does not allow these premises to be reconstructed in a unique way. It follows that deduction calls for an independent record of the premises, and such a record is provided by their propositional representations, which capture the propositions expressed by the premises. Some experimental evidence bears out the existence of such representations (see Mani & Johnson-Laird, 1982).

The model-building system, as the computer program shows, must be based on a set of underlying concepts. These *subconcepts* are built into the lexical semantics and the procedures for manipulating arrays, and they are based on increments to Cartesian coordinates, where the first coordinate is left–right, the second is front-back, and the third is up-down:

on the right of	1	0	0
on the left of	-1	0	0
in front of	0	1	0
behind	0	-1	0
above	0	0	1
below	0	0	-1

How the mind represents the meanings of spatial relations is not yet known, but it must deploy some set of underlying subconcepts that are used by the procedures for constructing models. These subconcepts are ineffable; they are not available to introspection, and they are probably innate and universal. They and the procedures that use them govern the mapping from linguistic expressions to models, and the mapping from models back to linguistic expressions again. Hence, the operations for reasoning by building mental models are essentially conceptual: reasoners use their understanding of descriptions to envisage situations, and this understanding ultimately depends on tacit conceptual knowledge. Readers should now be able to understand the essential characteristics of mental models. The key feature of a spatial model, for example, is not that it represents spatial relations, because a propositional representation can also do that, but its structure. In particular, the model is functionally organized in terms of axes so that information in it can be accessed by values on these axes. Such an organization in a mental model does not necessarily imply that information is laid out physically in a spatial way in the brain. It could be laid out in this way, but it need not be. The spatial reasoning program relies on arrays, which are a standard form of data structure in the programming language LISP, but these data structures are only functionally arrays and no corresponding physical arrays of data are likely to be found in the computer's memory. The same functional principle is likely to apply to high-level spatial models in human cognition.

2. Errors in Reasoning from Double Disjunctions

If the model theory is correct, then it should be possible to test human deductive competence to the point of breakdown merely by increasing the number of models. Johnson-Laird, Byrne, and Schaeken (1992) have confirmed this prediction using so-called "double disjunctions," such as:

1. Jane is in Seattle or Raphael is in Tacoma, but not both.
2. Jane is in Seattle or Paul is in Philadelphia, but not both.
3. What follows?

Each premise calls for two models, but their combination yields the following two models:

where "s" denotes Jane in Seattle, "t" denotes Raphael in Tacoma, and "p" denotes Paul in Philadelphia. The two models support the conclusion:

jane is in Seattle, or Raphael is in Tacoma and Paul is in Philadelphia.

If, instead of exclusive disjunctions, the premises are based on inclusive disjunctions:

1. Jane is in Seattle or Raphael is in Tacoma, or both.
2. Jane is in Seattle or Paul is in Philadelphia, or both.
3. What follows?

then they yield five alternative models:

s	t	p
s	t	
s		p
s		
	t	p

These models support the conclusion:

Jane is in Seattle, or Raphael is in Tacoma and Paul is in Philadelphia, or both.

The assertion that one of the two disjunctions is true and one is false calls for an exclusive disjunction of them, and the models for an exclusive disjunction, X or else Y, are according to Table 2:

X	
	Y

and so the disjunction calls for a list of all the models in the two alternatives. Hence, the problem as a whole calls for the following models:

k	
	a
k	a

If subjects judge the probabilities of two events by assessing which event occurs in more models, they will indeed infer that the ace is more probable than the king.

This conclusion is compelling, but it is wrong. If only one of the two assertions is true, then the other assertion is false: the two premises are in an exclusive disjunction, and so when one is true, the other is false. The models, however, represent only the true cases. When the false cases arc taken into account, the correct answer emerges. When the first disjunction is false there is *neither a king nor an ace,* and when the second disjunction is false there is *neither a queen nor an ace.* Either way, there is no ace—it cannot occur in the hand. Hence, the king, which can occur in the hand, is more probable than the ace, which cannot occur in the hand.

My colleagues and I have established that there are a number of different illusory inferences, which depend on a variety of sentential connectives. All that they appear to have in common is that their initial models, which fail to represent false contingencies, support conclusions that are quite wrong (i.e., that are contravened by fully explicit models of the premises). Thus, the third main prediction of the model theory is confirmed. Moreover, the existence of illusory inferences is contrary to all current theories based on formal rules of inference. They are based solely on valid rules of inference, and are accordingly unable to explain any phenomenon in which reasoners systematically draw invalid conclusions.

In general, experiments have confirmed all the main predictions of the model theory. These experiments cover the main domains of deduction, including syllogisms and inferences based on multiple quantifiers (see Johnson-Laird & Byrne, 1991).

IV. MODELS AND IMAGES

The mental model theory postulates that individuals represent verbal descriptions in the form of mental models, which are constructed from propositional represen-

As the theory predicts, the problems based on exclusive disjunctions (21% correct conclusions) were reliably easier than the problems based on inclusive disjunctions (8% correct conclusions). The problems were so difficult for the subjects, who were paid adult volunteers from all sorts of backgrounds, that they generally drew erroneous conclusions. The vast majority of these conclusions were based on only some of the models of the premises, typically just a single model. The results thus corroborated the second prediction of the model theory.

Other studies have shown that if reasoners reach a believable or congenial conclusion they tend not to search for alternative models (e.g., Oakhill, Johnson-Laird, & Garnham, 1989). This tendency is a frequent cause of everyday disasters, both minor and major. For example, the engineers in charge at Three-Mile Island inferred that a leak was the cause of the overheating of a relief valve, and overlooked the possibility that the valve was stuck open. The master of an English channel ferry, *The Herald of Free Enterprise,* inferred that the bow doors had been closed, and overlooked the possibility that they had been left open—an oversight that led to the drowning of several hundred passengers when the ferry capsized. The engineers at Chernobyl inferred that an explosion had damaged the reactor, and overlooked the possibility that there had been a meltdown of the reactor itself. The tendency to overlook possibilities seems an obvious danger. Yet strangely it cannot be predicted by formal rule theories, which have no elements within them corresponding to models of situations.

3. Illusory Inferences

Consider the following problem about a specific hand of cards:

One of the following assertions is true about the hand of cards, and the other assertion is false:

1. There is a king in the hand or there is an ace in the hand, or both.
2. There is a queen in the hand or there is an ace in the hand, or both.
3. Which is more likely to be in the hand: the king or the ace?

Recent studies carried out in collaboration with the author's colleagues, Fabien Savary and Patrizia Tabossi, have shown that almost all subjects tend to conclude:

The ace is more likely to be in the hand than the king.

The models of the first premise are:

k	
	a
k	a

and the models **of the second premise are:**

q	
	a
q	a

tations of the descriptions. Our task in Section IV is to elucidate the relations between models and images. We will show that a principled theoretical distinction should be drawn between them: models can contain elements that are not visualizable, and empirical evidence supports the existence of such elements in models. We then consider the sorts of mental operation that can be carried out on images, and how in many cases the transformation of an image may depend on an underlying model. This idea leads us to reconsider Kekulé's dream and to argue that problems cannot be solved by visual imagery alone.

A. How Models Differ from Images

Could the models that underlie reasoning be visual images? Some individuals report using imagery, but many do not—and their performance is equally predictable by the model theory. Indeed, we can go further: if reasoning depends on forming an image of the situation described in the premises, then an explicit manipulation of the "imageability" of the situation should affect inferential performance. A study carried out in collaboration with Ruth Byrne and Patrizia Tabossi examined this prediction (Johnson-Laird, Byrne, & Tabossi, 1989). The experiments used doubly quantified assertions, such as,

1. None of the artists is in the same place as any of the beekeepers.
2. All the beekeepers are in the same place as all the chemists.
3. What follows?

which can be represented in a model, such as:

| [a] [a] | [b] [b] [c] [c]

where a's denote artists, b's denote beekeepers, c's denote chemists, and the vertical bars demarcate separate places. This model supports the valid conclusion:

None of the artists is in the same place as any of the chemists.

We manipulated the "imageability" of the premises by using three sorts of relation: "taller than," "in the same place as," and "related to" (in the sense of kinship). An independent panel of judges rated how easy it was to visualize premises based on these relations, and their ratings differed significantly over the three relations. The experiment confirmed that one-model problems were easier than multiple-model problems, but there was no hint of an effect of imageability. Other experimenters have likewise failed to detect any influence of imageability on reasoning (see, e.g., Newstead, Manktelow, & Evans, 1982; Richardson, 1987). One can indeed reason from verbal descriptions that refer to abstract relations just as well as from those that refer to visualizable relations.

A major crisis occurred in the development of quantum mechanics, when the theory ceased to concern visualizable objects. Oddly, many mundane concepts are not visualizable either, but they are not in the least problematical in daily life. A good example is the ownership of property. One cannot perceive the relation between owner and owned, only *evidence* tor ownership. The concept of ownership can be glossed in the following terms (see Miller & Johnson-Laird, 1976, p. 558 et seq.):

> If an individual, x, owns an entity, y, then:
> i. it is permissible for x to use y, and it is not permissible for others to prevent this use.
> ii, it is permissible for someone else to use y it x gives permission ro this individual to do so; and it is permissible for x to give such permission.
> iii. x can act to transfer ownership of y to someone else, and such an action is permissible.

The details of this analysis are not so important as the general point: ownership hinges on deontic matters concerning what is permissible, and permissibility is not a visual property. One may visualize a permitted action; one may visualize a conventional symbol denoting permissibility (e.g., a check mark), but one cannot visualize the fact that an action is permissible.

An inferential system based on models can represent abstract elements. Earlier in the chapter, for example, we introduced the idea of negated elements in models, which are in essence "annotations" that serve a semantic function (Newell, 1990). Several studies have provided empirical evidence for such elements in models (Johnson-Laird & Byrne, 1991). Reasoners may of course use a visual image to represent negation (e.g., a large red cross that they superimpose on the model to be negated), but the image itself does not do the work of negation (as Wittgenstein, 1953, pointed out). The real work is done by the knowledge that the cross denotes negation, and by the system that uses this knowledge both to construct images and to interpret them. Reasoncrs do sometimes report using such images, but most people make no such reports and remain sublimely unaware of how they represent negation.

B. The Creative Manipulation of Images

The operations that are carried out in reasoning with models, as we argued earlier, are conceptual and semantic. But what about the operations that are carried out on images? What is their function? The question is complicated, but its answer will help to clarify the distinction between images and models. Images represent how something looks from a particular point of view—they may well be Marr's (1982) two-and-a-half-dimensional sketches, and operations on images are visual or spatial. Underlying an image of an object or scene is a three-dimensional (3-D) model, and operations on such models correspond to physical or spatial operations on the entities or scenes represented in the models. Hence, one way to explain the mental rotation experiments, which we described in Section II, is that visual system constructs a 3-D model from the first picture of the object, computes its major axis, and rotates the model via this axis to bring it into alignment with a model constructed from canonical shape, and this step presumably takes the extra amount of time. In another unpublished study, Richard Feit has implemented a computer program that will find an optimal match of three given shapes to a canonical representation. The algorithm is, technically speaking, intractable; that is, as the number of shapes to be matched increases so it takes exponentially longer to find a match. Human imagers arc also likely to be defeated by a large number of component shapes, and so it is an open question whether they, too, rely on an intractable procedure.

The "visual play" with images that occurs in these studies is reminiscent of scientists' reports of their own use of imagery in solving problems. Visualization alone, however, cannot solve problems. Just as models need to be backed up by propositional representations if they are to be used to reason, so too images need to be backed up by an independent representation of the problem if they are to be used to solve it. As an illustrative example, I return to Kekulé's problem, which I described at the start of the chapter, and reanalyze it in the light of Findlay's (1937) account.

Kekulé's goal was to formulate the molecular structure of benzene in terms of the theory of valencies—of which he was one of the founders. He knew that each benzene molecule contained six carbon atoms and six hydrogen atoms, and that the valency of carbon was four (i.e., each carbon atom should combine with four other atoms). The only known molecular structures at that time were in the form of strings, but a string of six carbon atoms required three hydrogen atoms to combine with each of the two atoms at its ends, and two hydrogen atoms to combine with each of the four atoms in the middle of the string. Hence, there did not seem to be enough hydrogen atoms to do the job. The puzzle stumped him until in mental play with a string-like structure, he formed an image of a circle. A merely circular arrangement of the atoms still does not solve the problem, because it calls for 12 hydrogen atoms. Kekulé had to make the further assumption that alternate links between the carbon atoms had a double valency, and so each carbon atom had a single link to one carbon atom and a double link to another. There remained only a single valency left unaccounted for, and it was the bond to a hydrogen atom. Because carbon atoms are identical, the single and double bonds oscillated from moment to moment. As Findlay (1937, p. 149) suggests, this oscillation may have been suggested by the atoms in his image "all twisting and twining in snake-like motion," but we do not know whether the solution really came to him in this way.

The manipulation of images is preeminently a method for solving visuospatial problems. Indeed, some theorists argue that the major function of visual imagery is to aid the process of object recognition: some perceptual cues trigger the synthesis of an image of an object from long-term knowledge, and the visual system tries to project this image onto the visual input (cf. Lowe, 1987; Marr, 1982). The manipulation of images accordingly yields spatial or physical rearrangements of entities. In solving problems, such as the structure of benzene, forming an image is only part of the process. The image must relate to an independent representation of the problem—just as the use of models in reasoning must relate to an independent the second picture. It is not the image that is rotated by an underlying model constructed from it. The evidence for this claim, and for the relative unimportance of the features of the 2-D pictures, is that rotations in depth produced the same pattern of results as rotations in the picture plane. As Metzler and Shepard (1982) remark:

> These results seem to be consistent with the notion that . . . subjects were performing their mental operations upon internal representations that were more analogous to three-dimensional objects portrayed in the two-dimensional pictures than to the two-dimensional pictures actually presented. (p. 45)

Operations on images' per se correspond to visual rearrangements. They can lead to the construction of new objects out of existing elements or shapes. Here, for example, is a task that the reader can carry out:

> Imagine the letter "B". Rotate it 90 degrees counter-clockwise. Put a triangle below it having the same width and pointing downwards. Remove the horizontal line. What have you got?

As Finke, Pinker, and Farah (1989) have shown, individuals can carry out such tasks with reasonable success, and they are unable to predict the outcome—they have to carry out the operations on their images in order to "see" the result. The answer in the present case is a heart shape.

Finke and his colleagues have also shown how the manipulation of images can yield a creative result. They used an array of the following 15 simple shapes: circle, square, triangle, rectangle, horizontal line, D, I, L, T, C, J, 8, X, V, P, (see Finke & Slayton, 1988). On each trial, the experimenter named three of the shapes, and the subjects had to close their eyes and to imagine assembling these shapes into a recognizable figure. The subjects were free to combine the parts in any way: rotating, translating, superimposing, or juxtaposing them. They could change their sizes, but they were not allowed to distort the shapes. They were surprisingly successful at this task, and about 15% of their efforts were rated by an independent panel of judges as creative. The task seems to depend on a creative "play" with images, and again the subjects were wholly unable to predict in advance what its outcome would be.

In an unpublished study, Jung-Min Lee has observed that a constraint on the required outcome can speed up the process. The subjects had to use any three shapes from Finke s array to synthesize an image. They were reliably faster to synthesize a so-called basic-level object (Rosch, 1977), such as a house, an apple, or a chair, than an instance of a superordinate category, such as a building, a fruit, or a piece of furniture. Instances of a basic-level object, as Rosch showed, have more uniform shapes than instances of their superordinates (e.g., two chairs are likely to be more similar in shape than two pieces of furniture). Hence, the subjects could proceed by imagining the canonical shape of, say, a chair, search the array for appropriate elements, such as the "L" shape and the square, and then assemble them. With superordinates, this strategy can be used only after the subject has called to mind an object with a representation of the premises. Visualization can yield deep conceptual innovations only within a system that can represent more abstract information.

C. Images and Diagrams

The moral of the previous section may lead skeptics to dismiss imagery as epiphenomenal: the real work in conceptual innovation is done, they may say, by underlying propositional representations. In this section, I intend to rebut this argument on the basis of experimental evidence, and to show that images cannot be reduced to propositional representations; both are high-level representations necessary to explain thinking. The argument will hinge on the role of diagrams in reasoning.

Diagrams are often said to be helpful aids to thinking. They can make it easier to find relevant information—one can scan from one element to another element nearby much more rapidly than one might be able to find the equivalent information in a list of numbers or verbal assertions (see also Cutting & Massivoni, chap. 6, this volume). Diagrams can make it easier to identify instances of a concept—an iconic representation can be recognized faster than a verbal description. Their symmetries can cut down on the number of cases that need to be examined. But can diagrams help the process of thought itself? Larkin and Simon (1987) allow that diagrams help reasoners to find information and to recognize it, but doubt whether they help the process of inference itself. Barwise and Etchemendy (1992), who have developed a computer program, Hyperproof, for learning logic, write:

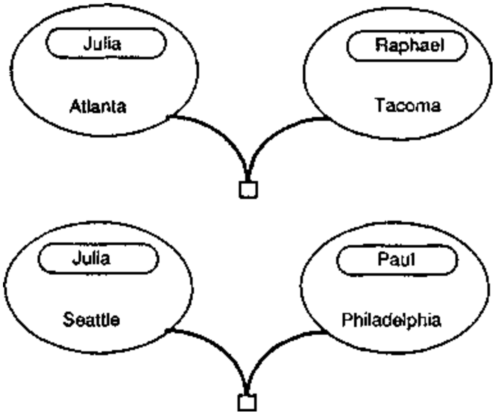
> Diagrams and pictures are extremely good at presenting a wealth of specific, conjunctive information. It is much harder to use them to present indefinite information, negative information, or disjunctive information. For these, sentences are often better. (p. 82)

Hyperproof accordingly captures conjunctions in diagrams, but expresses disjunctions in verbal statements. The model theory, however, makes a different prediction. The problem in reasoning is to keep track of the possible models of premises. Hence, a diagram that helps to make them explicit should also help people to reason.

Malcolm Bauer and the author tested this prediction in two experiments based on double disjunctions (Bauer & Johnson-Laird, 1993). In the first experiment, the premises were either in a verbal form, such as:

1. Julia is in Atlanta or Raphael is in Tacoma, or both.
2. Julia is in Seattle or Paul is in Philadelphia, or both.
3. What Mows?

or else in the form of a diagram, such as Figure 1. To represent, say, Julia in Atlanta, the diagram has a lozenge labeled "Julia" lying within the ellipse labeled "Atlanta." Inclusive disjunction, as the figure shows, is represented by a box connected by lines to the two component diagrams making up the premise as a whole. The experi-



What follows?

FIGURE 1 The diagram representing a double disjunction (a negative inclusive one) in the first diagram experiment.

ment confirmed that exclusive disjunctions were easier than inclusive disjunctions (for both the percentages of correct responses and their latencies); it also confirmed that problems in which the individual common to both premises was in the *same* place in both of them ("affirmative" problems) were easier than problems in which the individual common to both premises was in different places in them ("negative" problems, such as the one above). But the experiment failed to detect any effect of diagrams: They yielded 28% correct conclusions in comparison to the 30% correct for the verbal problems. Double disjunctions remained difficult and diagrams were no help at all.

With hindsight, the problem with the diagrams was that they used arbitrary symbols to represent disjunction and thus failed to make the alternative possibilities explicit. In a second experiment, we used a new sort of diagram, as shown in Figure 2, analogous to an electrical circuit. The idea, which we explained to the subjects, was to complete a path from one side of the diagram to the other by moving the shapes corresponding to people into the slots corresponding to cities. We tested four separate groups of subjects with logically equivalent problems: one group received diagrams of people and places (as in Figure 2); one group received problems in the form of circuit diagrams ot electrical switches; one group received problems in the form of verbal premises about people and places, and one group received problems in the form of verbal premises about electrical switches. There was no effect of the content of the problems—whether they were about people or switches—and so we have pooled the results. The percentages of correct responses are presented in Figure 3. As the figure shows, there was a striking effect of mode

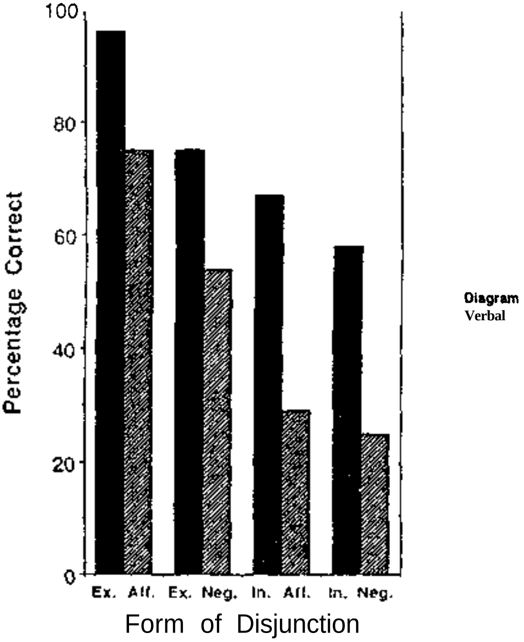


FIGURE 3 The percentages of correct conclusions in the second diagram experiment. There are four sorts of problem: Ex. Aff., affirmative problems based on exclusive disjunctions; Ex. Neg., negative problems based on exclusive disjunctions; In. Aff., affirmative problems based on inclusive disjunctions; In. Neg., negative problems based on inclusive disjunctions.

V. CONCLUSIONS

This chapter has focused on thinking with a propositional content. Other sorts of thinking lack such content, such as the mental processes controlling the improvisations of a musician, the gestures of a painter, or the movements of a dancer. The aim of the chapter was to understand the role of visualization in thinking about propositions, and it has reached three main conclusions.

First, thinking depends **on** propositional representations and mental models. Propositional representations capture the meaning of premises, and they are used to construct mental models representing the situation under discussion. Reasoners test the strength of an inference (or argument) by searching for alternative models of the propositional representations in which a putative conclusion is false. The experimental evidence corroborates the theory's predictions: more models mean more work, erroneous conclusions are a result of overlooking possible models, and illusory inferences arise from a failure to represent false contingencies.

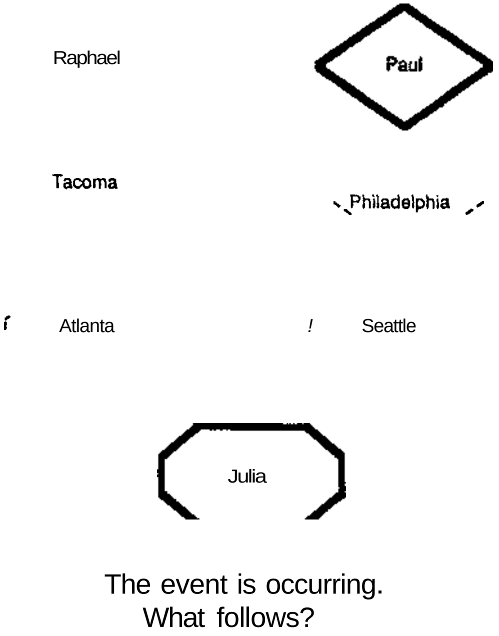


FIGURE 2 The diagram representing a double disjunction (a negative inclusive one) in the second diagram experiment.

of presentation: 74% correct responses to the diagrammatic problems in comparison to only 46% correct responses to the verbal problems. The results also corroborated the model theory's predictions that exclusive disjunctions should be easier than inclusive disjunctions, and that affirmative problems should be easier than negative problems. The latencies of the subjects' correct responses had exactly the same pattern, and they were reliably faster to respond to the diagrammatic problems (mean of 99 sec) than to the verbal problems (mean of 135 sec).

People evidently reason by trying to construct models of the alternative possibilities, and diagrams that enable these alternatives to be made explicit can be very helpful. In a series of unpublished studies, Victoria Bell and the author have found that merely teaching people to maintain lists of the separate possibilities, either in diagrams or in the mind's eyes, improves their reasoning. Likewise, with a diagram of the sort shown in Figure 2, individuals perceive the layout and in their mind's eye they can move people into places and out again. By manipulating a visual image that has the external support of a corresponding diagram, they can construct the alternative possibilities more readily than they can do so from verbal descriptions. It follows that diagrams are not merely encoded in propositional representations equivalent to those constructed from descriptions.

Second, models and images differ. Models can be 3-D, and can embody abstract predicates that are not visualizable. Hence, they can represent any situation, and operations on them can be purely conceptual. In contrast, images represent how something looks from a particular point of view. They are projected from the visualizable aspects of underlying models. Images and diagrams, however, can be used in a symbolic way. If one wishes to convey what is going on in a complex domain with many varying numerical quantities, such as the flow of air around an airplane, then the translation of the data into a visual display can capitalize on the power of the visual system to extract high-level patterns from low-level data. To make sense of an array of 100 million numbers (the intensities of light falling on the cells in the retina) the brain has "software" that uses these data to construct a high-level model of the world suitable for the limited powers of consciousness. The visual display is symbolic; that is, it does not correspond directly to the external world. Our experimental study with diagrams showed that most individuals are able to imagine moving a shape from one position to another and in this way to envisage a proposition with a wholly different content (people in places). Some logicians claim that diagrammatic methods of reasoning are in some way improper (cf. Tennant, 1986), but Barwise and his colleagues have shown that they are valid, and indeed can be complete systems, which capture all valid inferences (Barwise & Etchemendy, 1991; Shin, 1992).

Third, people can construct novel images out of given components. They can retrieve the canonical shape of an object and then in their mind's eye assemble that shape out of the preexisting components—a process that calls for moving one shape in relation to another, juxtaposing or superimposing them, and so on. Humans can simulate phenomena dynamically, and some individuals spontaneously carry out such simulations (e.g., Nicola Tesla, the inventor, was said to be able to imagine the wear in his machines by simulating running them in his mind's eye; Shepard, 1978).

Could visualization lead to profound innovations and novel scientific concepts? Could it, for example, lead from an Aristotelian concept of velocity to the Newtonian concept of instantaneous velocity, or from an absolute concept of simultaneity to a relativistic concept? The answer in our view is that it could play a part in such transitions, but that lying behind a scientist's "picture" of the world is likely to be a mental model representing more abstract relations, and its associated subconceptual apparatus. Visualization can help thinkers to envisage possibilities, and it may help them to imagine certain spatial and physical properties and operations. They cannot, however, directly visualize abstract concepts or conceptual relations. Manipulations of an image can be reinterpreted in terms of the model and can lead to conceptual innovations. Kekulé's visual manipulations of snake-like images is one such example. Thinking also depends on more than models or images. They can be exploited only within a system that carries out conceptual operations. The underlying machinery depends on a set of subconceptual elements, which are tacit, primitive, and probably innate. The interplay between models and subconcepts is the most likely locus of conceptual innovation. It depends on a new sort of model, which typically embodies concepts that are neither observable nor visualizable, and in turn rest on the construction of new concepts from the subconceptual repertoire. Such models may generate fresh problems and, as Wise (1979) points out, their solution may call for a reorganization of the models themselves.

References

Baddeley, A. E., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol. 8* (pp. 278-301). London: Academic Press.

Barwise, J., & Etchemendy, J. (1991). Visual information and valid reasoning. In W. Zimmermann & S. Cunningham (Eds.), *Visualization in teaching and learning mathematics* (pp. 120–132). Mathematical Association of America.

Barwise, J., & Etchemendy, J. (1992). Hyperproof: Logical reasoning with diagrams. In N. H. Narayanan (Ed.), *AAAI Symposium on reasoning with diagrammatic representations* (pp. 80-84). Stanford, CA: AAAI.

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science, 4*, 372-378.

Binet, A. (1894). *Psychologie des grands calculateurs et joueurs d'echecs.* Paris: Hachette.

Boltzmann, L. (1974). On the significance of physical theories. In. B. McGuiness (Ed.), *Ludwig Boltzmann: Theoretical physics and philosophical problems* (pp. 33–36). Boston: Reidel. (Original work published 1890)

Boltzmann, L. (1974). On the fundamental principles and equations of mechanics. In B. McGuiness (Ed.), *Ludwig Boltzmann: Theoretical physics and philosophical problems* (pp. 101–128). Boston: Reidel. (Original work published 1899)

Bower, G. H. (1970). Analysis of the mnemonic device. *American Scientist, 58*, 496-501.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation, Vol. 18.*, (pp. 313-371). New York: Academic Press.

Byrne, R. M. J., & Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory and Language, 28*, 564-575.

Craik, K. (1943). *The nature of explanation.* Cambridge: Cambridge University Press.

Darwin, C. (1859/1968). ***The origin of species by means of natural selection** or the preservation of favoured races **in the** struggle for life.* **J. W. Burrow** (Ed.). **New York: Viking Penguin.**

Feldman, D. H. (1988). Dreams, insights, and transformations. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 271-297). Cambridge, UK: Cambridge University Press.

Ferguson, E. S. (1977). The mind's eye: Nonverbal thought in technology. *Science, 197*, 827–836.

Findlay, A. (1937). *A hundred years of chemistry.* London: Duckworth.

Finke, R. A., Pinker, S., & Farah, M. (1989). Reinterpreting visual patterns in mental imagery. *Cognitive Science, 13*, 51-78.

Finke, R. A., & Slayton, K. (1988). Explorations of creative visual synthesis in mental imagery. *Memory and Cognition, 16*, 252-257.

Galton, F. (1928). *Inquiries into human faculty and its development.* London: Dent. (Original work published 1880)

Gleick, J. (1992). *Genius: The life and science of Richard Feynman.* New York: Pantheon Books.

Gruber, H. E. (1974). *Darwin on man: A psychological study of scientific creativity.* (2nd ed.). Chicago: University of Chicago Press.

Gruber, H. E. (1994). Insight and affect in the history of science. In R. J. Sternberg and J. Davidson (Eds.), *The nature of insight* (pp. 397-431). Cambridge, MA: MIT Press.

Hadamard, J. (1996). *The Mathematician's Mind: The psychology of invention in the mathematical field.* Princeton: Princeton University Press. (Original work published 1945).

Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996). Mental models and temporal reasoning. *Cognition, 60,* 205-234.

Shepard, R. N. (1978). Externalization of mental images and the act of creation. In B. S. Randhawas & W. E. Coffman (Eds.), *Visuallearning, thinking, and communication* (pp. 133-189). New York: Academic Press.

Shepard, R. N. (1988). The imagination of the scientist. In K. Egan and D. Nadaner (Eds.), *Imagination and Education.* New York: Teachers College Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701-703.

Shin, Sun-Joo (1992). A semantic analysis of inference involving Venn diagrams. In N. H. Narayanan (Ed.), *AAAI Symposium on reasoning with diagrammatic representations* (pp. 85-90). Stanford, CA: AAAI.

Störring, G. (1908). Experimentelle Untersuchungen über einfache Schlussprozesse. *Archiv fürdie gesamte Psychologie, 11,* 1-27.

Tennant, N. (1986). The withering away of formal semantics. *Mind and Language, 1,* 302–318.

Valéry, P. (1894/1972). Introduction to the method of Leonardo da Vinci. In *Leonardo Poe Mallarmé. Vol. 8 ofthe Collected Works ofPaul Valéry* (pp. 3-63). Princeton, NJ: Princeton University Press.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology ofreasoning: Structure and content,* Cambridge, MA: Harvard University Press.

Wertheimer, M. (1961). *Productive thinking.* M. Wertheimer (Ed.). London: Tavistock.

Wise, M. N. (1979). The mutual embrace of electricity and magnetism. *Science, 203,* 1310-1318.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: Macmillan.

Wotiz, J. H., & Rodofsky, S. (1984). Kekulé's dreams: Fact or fiction? *Chemistry in Britain,* 720-723.

Hagert, G. (1984). Modeling mental models: experiments in cognitive modeling of spatial reasoning. In T. O'Shea (Ed.), *Advances in artificial intelligence.* (pp. 121-134). Amsterdam: North-Holland.

Holt, R, R. (1964). Imagery: The return of the ostracized. *American Psychologist, 19,* 254–264.

Humphrey, G. (1951). *Thinking.* London: Methuen.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1994). Mental models and probabilistic thinking. *Cognition, 50,* 189-209.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N., Byrne, R. M.J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99,* 418-439.

Johnson-Laird, P. N., Byrne, R. M. J., & Tabossi, P. (1989). Reasoning by model: The case of multiple quantification. *Psychological Review, 96,* 658-673.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kosslyn, S. M., Ball, T. M., & Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 47–60.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth 10,000 words. *Cognitive Science, 11,* 65–99.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31,* 355-395.

Luria, A. R. (1969). *The mind of a mnemonist.* London: Cape.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory and Cognition, W,* 181-187.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco: W. H. Freeman.

Maxwell, J. Clerk (1911). Diagram. In *The Encyclopaedia Britannica. Vol. XVIII.* New York: The Encyclopaedia Britannica Company.

Metzler, J., & Shepard, R. N. (1982). Transformational studies of the internal representations of three-dimensional objects. In R. N. Shepard & L. A. Cooper, *Mental images and their transformations* (pp. 25–71). Cambridge, MA: MIT Press. (Original work published in Solso, R. L. (Ed.) *Theories in Cognitive Psychology: The Loyola Symposium.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1974.)

Miller, A. (1984). *Imagery in scientific thought: Creating 20th-century physics.* Boston, MA: Birkhauser.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Newstead, S. E., Manktelow, K. I., & Evans, J. St. B. T. (1982). The role of imagery in the representation of linear orderings. *Current Psychological Research, 2,* 21-32.

Oakhill, J. V., Johnson-Laird, P. N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition, 31,* 117-140.

Ohlsson, S. (1984). Induced strategy shifts in spatial reasoning. *Acta Psychologica, 57,* 46-67.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart, and Winston.

Perky, C. W. (1910). An experimental study of imagination. *American Journal of Psychology, 21,* 422-452.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80,* 1-24.

Richardson, J. T. E. (1987). The role of mental imagery in models of transitive inference. *British Journal of Psychology, 78,* 189-203.

Rips, L. J. (1994). *The psychology of proof.* Cambridge, MA: MIT Press.

Rosch, E. (1977). Classification of real-world objects: Origins and representations in cognition. In P. N. Johnson-Laird & P. C.Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 212-222). Cambridge: Cambridge University Press.