cleverly) characterizing it in terms of other already given mathematical objects.

2 For further discussions, rich with examples, see Woodward (1989), Harper (1990), and Kaiser (1991).

3 By 'abstract' I mean not in space and time; numbers, properties, and propositions are typical abstract entities. Nominalists take such entities to be mere words; realists take them to be genuine objects existing independently of humans. My view of phenomena is, of course, a realist account.

4 For more on thought experiments, see my *The Laboratory of the Mind: Thought Experiments in the Natural Sciences* (1991).

5 But not exactly momentum, since Descartes eschewed mass. For him 'quantity of motion' would be more like 'size of matter times speed.' Leibniz did not have a clear notion of mass either, though unlike Descartes, he was not wedded to a purely kinematic physics.

6 This challenge came from Simon Blackburn.

7 For more on natural kind reasoning, see Harper (1989). Generally, this is unexplored territory and deserves a great deal more attention.

8 Thanks to Mary Tiles for this point. I'm grateful to her for helpful discussions on a number of other topics as well.

9 The themes of this paper are treated in my book *Smoke and Mirrors: How Science Reflects Reality* (1994). I am also very glad to acknowledge the financial help of the Social Sciences and Humanities Research Council of Canada, and finally I wish to thank David Kotchan for drawing several of the diagrams.

9. Visual Models and Scientific Judgment

RONALD N. GIERE

1. INTRODUCTION

When reading scientific papers or watching presentations by scientists, nothing is more obvious than the use of *visual* modes of presentation for both theory and data. This not a new phenomenon, although it has been emphasized recently by the development of computer graphics. One finds a widespread use of various visual devices going back to the Scientific Revolution. Newton's *Principia*, for example, is full of diagrams used in his geometrical demonstrations. But why should anyone be particularly interested in the use of pictures and diagrams in science? Specifically, why should a *philosopher of science* be interested in this particular aspect of the practice of science?

It is my view that studying visual modes of representation in science provides an entrée into fundamental debates within the philosophy of science, as well as in related fields such as the history, psychology, and sociology of science. I will begin by indicating the nature of these debates and pointing out the relevance to these broader issues of the role played in science by visual modes of presentation. In the latter part of the paper, I will use some diagrams that played a central role in the twentieth-century revolution in geology in order to illuminate these general themes.[1]

2. GENERAL ISSUES

Within the English-speaking world, the logical empiricist image of science, and the projects it generated, dominated philosophical thought about science for a generation following the Second World War. Two fundamental aspects of this image are relevant here. First, scientific knowledge consists primarily of what is encapsulated in scientific theories, and theories are ideally to be thought of as interpreted axiomatic systems. It follows that the primary mode of representation in science is *linguistic* representation. Second, the reasoning which legitimates the claims of a particular theory as genuine knowledge has the general character of a logic. That is, there are rules which operate on linguistic entities yielding a 'conclusion' or some other linguistic entity such as a probability assignment.

In the framework of logical empiricism, then, there can be no fundamental role in science for non-linguistic entities like pictures or diagrams. Such things might, of course, play some part in how scientists actually learn or think about particular theories, but unless their content is reduced to linguistic form, they cannot appear in a *philosophical* analysis of the content or legitimacy of any scientific claims to knowledge.

Like so many other aspects of post–Second World War Western culture, the logical empiricists' picture of science began to blur in the decade of the 1960s. A major stimulus for change, and focus for opposing views, was Thomas Kuhn's *Structure of Scientific Revolutions* (1962). The initial rejection of Kuhn's views by philosophers of science was to be expected because he rejected the major assumptions of logical empiricism. According to Kuhn, for example, general statements organized into axiomatic systems play little role in the actual practice of science. There is thus little to be learned about science by reconstructing theories in a logical empiricist mould. Moreover, the relative evaluation of rival paradigms is not something that can be reduced to any sort of logic. It is fundamentally a matter of choice by scientists acting as individuals within a scientific community. For Kuhn, science is primarily a puzzle-solving activity. Scientific revolutions are the result of many individual scientists making the judgment that a particular type of puzzle, or way of approaching puzzles, is no longer fruitful, and that another approach provides a more promising basis for further puzzle-solving activities.

Kuhn himself did not highlight the role of visual or other non-propositional modes of representation in science. Indeed, he avoided talk about representation. I surmise that was largely because he, like most everyone else, thought of representation in propositional terms, and that leads immediately to the concept of truth. His picture of science as a puzzle-solving activity was meant to be an alternative to the view of science as producing truths. Moreover, his emphasis on the incommensurability of terms in the languages of rival paradigms shows his tendency to think of science in linguistic categories.[2]

Nevertheless, Kuhn's approach to understanding science at least opened the door to consideration of non-linguistic representational devices in the practice of science. This was not just because his account was historical, but because it was *naturalistic*. He was trying to explain how science works in terms of naturalistic categories like the psychological make-up of individual scientists and the social interactions among scientists in communities. Thus, whether non-propositional devices like diagrams and graphs play a significant role in science is something to be determined empirically by examining actual cases of science in action.

Philosophers were initially quick to charge Kuhn with having fallen into epistemological relativism, a charge he personally has struggled to avoid.[3] But beginning in the mid-1970s, several groups of European *sociologists* of science have pushed the relativistic aspects of Kuhn's views to their logical conclusion. The slogan of these schools is that science is a *social construct*. The import of the slogan is most quickly grasped by reflecting on the extent to which *society* is a social construct. There is, for example, nothing in the non-human universe that requires representative democracy, an independent judiciary, separation of church and state, or any other of the fundamental structures of American society. These are historically conditioned social constructs. Science, it is claimed, is no different. It follows that the world-view of those we call 'primitive' is in no objective way inferior to ours. It is just different. The only thing special about our scientific world-view is that it is ours.[4]

Significantly, relativist sociologists of science were among the first to investigate the role of pictures, diagrams, and other non-propositional forms of representation in science. Their aim has been to show how images are created and deployed in the social construction of scientific knowledge. The initially plausible view that these various images somehow picture reality is thereby 'deconstructed.'

There are more radical and less radical strains within the constructivist camp. A less radical view is to admit that scientists intend their theories to represent the world and often believe that they have succeeded. It is just that close sociological and anthropological analysis reveals that the intentions are not fulfilled and the beliefs mistaken. A more radical view is that science is not really a representational activity after all. In the twentieth century, painting clearly moved from being essentially representational to allowing forms that are not representational at all. We now have pictures that are not pictures of anything, and were never intended to be. So, it might be claimed, science is now (and maybe always has been) non-representational. Our theories don't picture anything.[5]

My view is that what is needed is a middle way between philosophical positivism and sociological relativism, both of which, in very different ways, deny any genuine representational role for visual images in science. Examining visual modes of theorizing and evaluating data is part of a strategy for developing the desired middle way. Since images could not literally be true or false, this strategy avoids raising questions about the nature of truth. It thus makes possible the pursuit of a naturalistic theory of science which goes beyond puzzle solving to explore ways in which visual models might genuinely represent the real world, and be correctly judged to do so.

As just indicated, there are several major parts to the overall program of developing a naturalistic middle way. A major task, for example, is simply to understand the various ways images and other non-propositional devices can be used to represent the world. Here I will approach this task only to the extent of pointing out how a model-based understanding of scientific theories makes it possible to treat things like diagrams and scale models on a par with the more abstract theoretical models that, on this account, form the core of any scientific theory. The focus of this paper will be on explaining how pictorial presentations of data can be used in judging the relative representational adequacy of visually presented models of the world. Or, to put it in more traditional terms, I want to present (part of) a theory of scientific reasoning in which visual presentations of both data and theory can play a significant role. The 1960s revolution in geology provides a particularly rich context for just such a presentation.

3. MODELS AND THEORIES

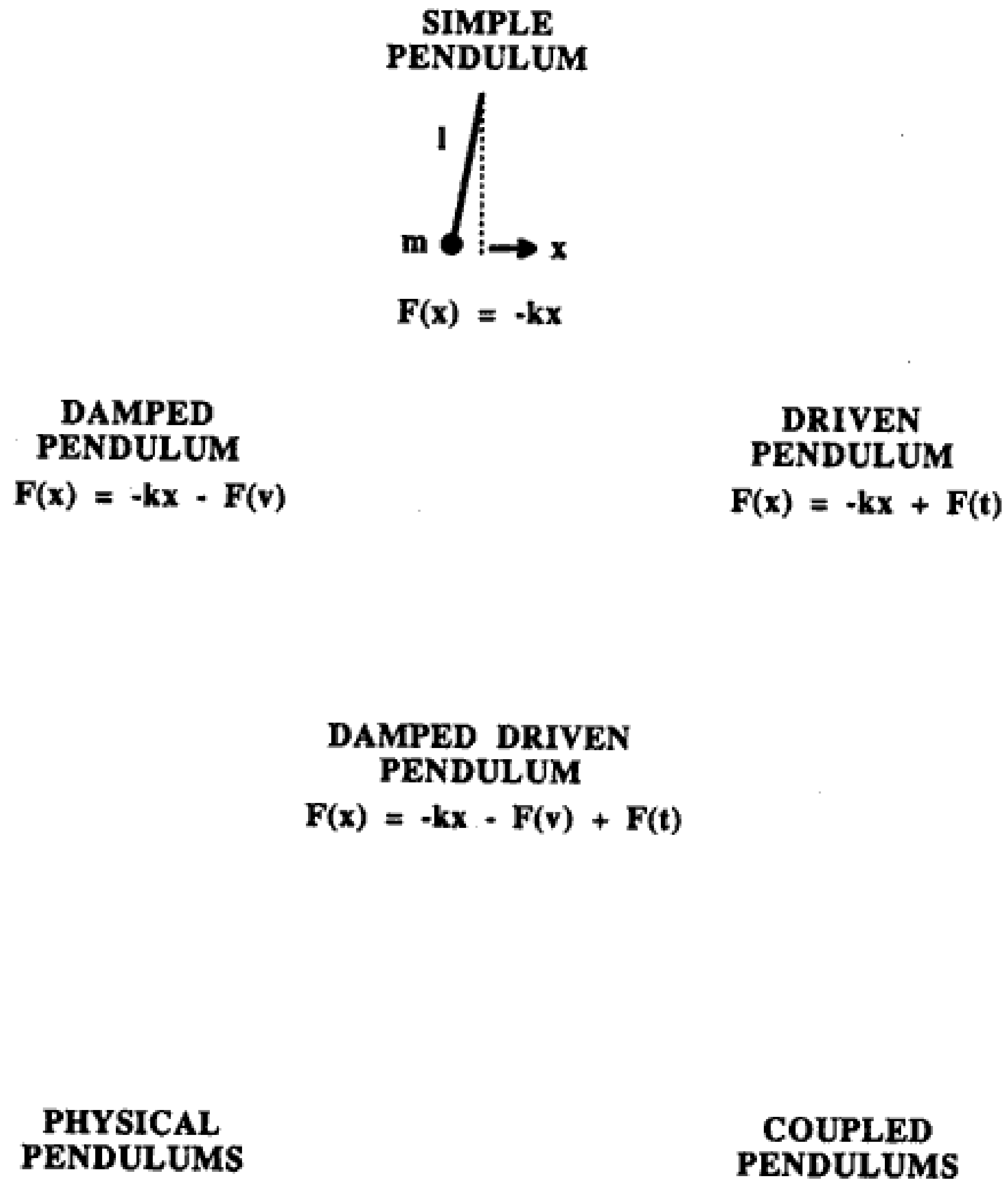
For a generation now, a number of philosophers of science have been developing an alternative to the logical empiricist account of scientific theories. This account has several names. It is sometimes called the 'semantic view of theories,' by way of contrast with the supposed 'syntactic' character of theories on the received view. It is also called

the 'non-statement' view, the 'predicate' view, or (as I now prefer) the 'model-based' view of theories.[6]

A common way of describing the model-based view is to say that theories include two different sorts of linguistic entities. Some are predicates, which may have a quite elaborate internal structure, as, for example, the predicates 'pendulum' or 'two-body Newtonian gravitational system.' Others are statements of the form 'X is P' where X refers to a real-world system and P is one of the predicates, as in the statement 'The earth-moon system is a two-body Newtonian gravitational system.' The predicates, as such, have no truth values, but the associated statements do.

This way of characterizing the statement and model-based views of scientific theories makes the differences between them seem relatively formal, even trivial. Expressions that function as empirical laws on the statement view, for example, reappear in the definitions of predicates on the model-based view. And there seem to be few if any significant empirical claims that could be formulated in one framework and not in the other. In the present context, however, the main difficulty with this way of formulating the difference is that it overemphasizes the *linguistic* aspects of the model-based approach. A way to redress this deficiency is to shift one's focus away from the predicates to the objects they encompass.

On my understanding of a model-based approach to scientific theories, the predicate 'pendulum,' as it appears in classical mechanics, does not apply directly to real-world objects like the swinging weight in the grandfather clock that stands in my living-room. It applies, rather, to a family of idealized models, the central example of which is the so-called 'simple pendulum.' A simple pendulum is a mass swinging from a massless string attached to a frictionless pivot, subject to a uniform gravitational force, and in an environment with no resistance. This is clearly an ideal object. No real pendulum exactly satisfies any of these conditions. So no real pendulum is a simple pendulum as characterized in classical mechanics. And the same is true for more complex types of classical pendulums: damped pendulums, driven pendulums, and so on. Figure 9.1 shows a family of models of pendulums radiating out from the model of a simple pendulum. So what is the relationship between the idealized model pendulums of classical mechanics and real swinging weights? It is, I suggest, like the relationship between a *prototype* and things judged sufficiently similar to the prototype to be classified as of



9.1 A family of models of pendulums radiating out from the model of a simple pendulum.

that type. And how are such judgments made? After all, any two objects (idealized or not) are similar to each other in infinitely many ways. Which features count for judgments of similarity to the prototype, and why do some features count more than others? Here there are no simple answers.

To some extent the models themselves provide guidelines for the relevant similarity judgments. The main dynamical variable in any model of a pendulum is the period of oscillation. It is a characteristic of the models that the mass of the bob is irrelevant to the period. Only

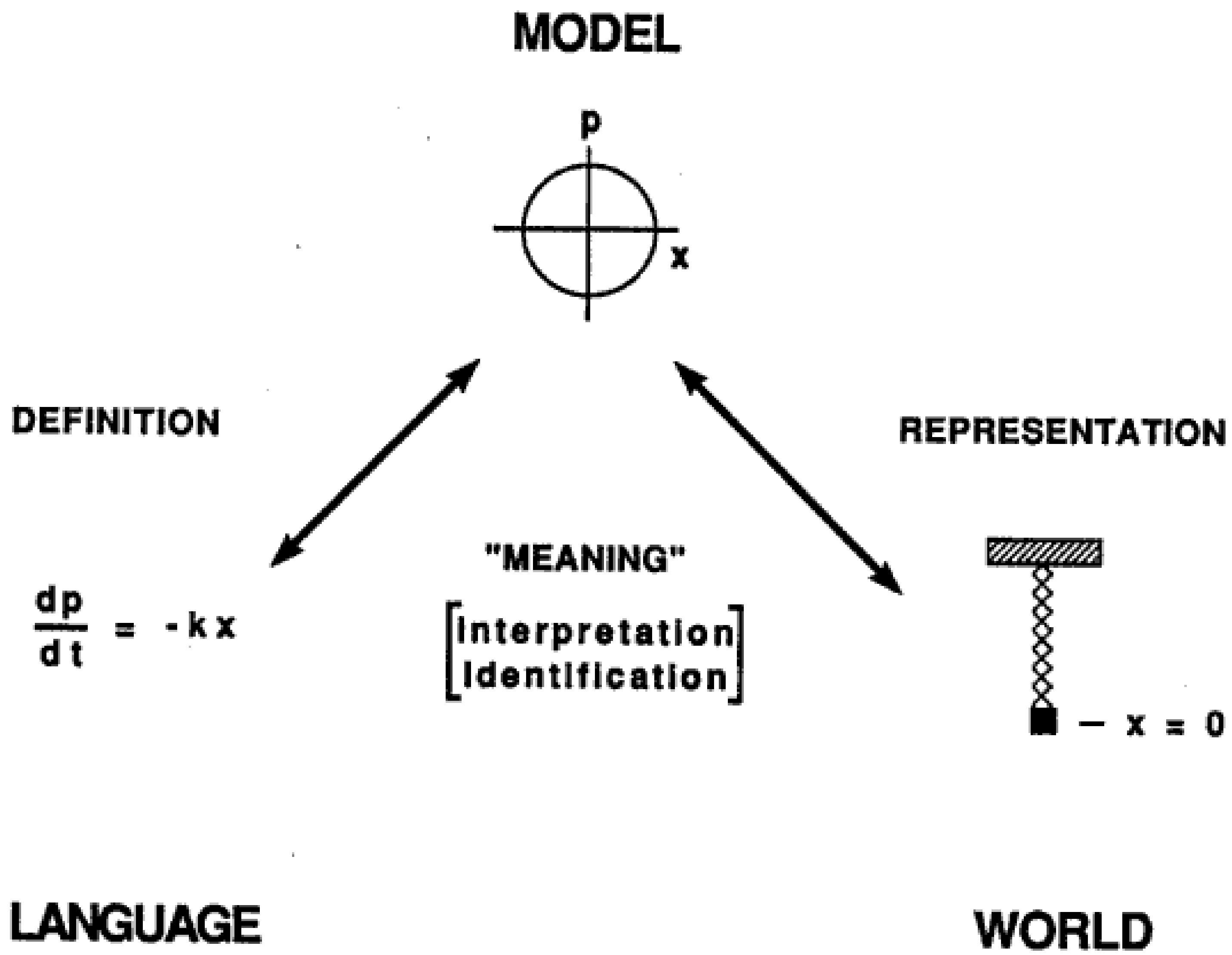
the length of the suspension, plus the gravitational force, matters. So the mass of the bob should be relatively unimportant in classifying some real swinging weight as a simple pendulum. So should its shape. Yet if one is building an ordinary grandfather clock, a one-pound pie-shaped bob swinging in air and a one-ton spherical bob swinging in water will not be regarded as equally appropriate approximations to a simple pendulum, even though the deviation in period from the ideal might be similar. Other, highly practical, considerations are overriding.

Figure 9.2 is an attempt to picture the relationships among representational devices such as language, models, and objects in the real world. Important and interesting though these relationships may be, I cannot further pursue these general issues here.[7] The main point for present concerns is that, on this view of scientific theories, the primary representational relationship is not the truth of a statement relative to the facts, or even the applicability of a predicate to an object, but the similarity of a prototype to putative instances. This is not a relationship between a linguistic and a non-linguistic entity, but between two non-linguistic entities. Once this step has been taken, the way is clear to invoke other, less abstract, non-linguistic entities to play a similar role.

Consider the sketch of a simple pendulum shown at the top of figure 9.1. I would regard this diagram as a particular embodiment of the abstract model of a simple pendulum. It too can serve as a prototype in judging the similarity of a particular real pendulum to the classical model of a simple pendulum. What holds for this simple diagram should in principle apply to a host of other non-linguistic representational devices. The difficulties in getting from 'in principle' to 'in practice' should not be underestimated, but they are not my main concern here.

4. CRUCIAL DECISIONS

The idea of a 'crucial experiment,' as expounded, for example, by Francis Bacon, was a major cultural achievement of the Scientific Revolution. It deserves to be ranked along with such other achievements as the calculus, the telescope, and the air pump. How it came to have that status is a difficult historical question. Of course, with the hindsight of three centuries, we know that the role of crucial experiments has often been exaggerated, and that the designation of an experiment as 'crucial' often comes long after the fact. But this only shows that the idea of a crucial experiment can play a rhetorical as well



9.2 Relationships among language, models, and objects in the real world.

as an operational role in science. It does not show that, properly understood, it plays no operational role.[8]

But what is the 'proper' operational understanding of crucial experiments? Here there are as many answers as there are approaches to the general problem of theory evaluation in science. For example, on a purely *deductive* account of scientific reasoning, the 'logic' of crucial experiments is just the logic of disjunctive syllogism plus modus tollens. Suppose T_1 and T_2 form an exclusive and exhaustive disjunction. And suppose that T_1 implies *O* while T_2 implies *Not-O*. The 'crucial experiment' yields the 'observation' *O*. By modus tollens, T_2 is falsified, thus justifying T_1 by disjunctive syllogism.

Again, on a *probabilistic* account of theory evaluation, we suppose that T_1 and T_2 have comparable initial probabilities. A 'crucial experiment' would be one for which the final probability of T_1 given the observed outcome is much greater than the final probability of T_2. Recently it has been argued that theory evaluation is primarily a matter of the relative *explanatory coherence* of the rival theories with given observations. Here

the explanatory coherence of a theory is a function of coherence relationships among statements.[9]

There is no need here to rehearse the many reasons that might be given for rejecting these approaches to scientific reasoning.[10] I wish only to point out that they all assume a *propositional* account of scientific theories. It is statements that are falsified, assigned low probabilities, or cohere. There is therefore no way that visual or other non-propositional forms of information can play a role in the reasoning without first being reduced to propositional form. I will now outline an account of crucial experiments that allows reasoning based directly on visual images. It assumes a model-based understanding of scientific theories along the lines outlined above.

The label 'crucial decisions' already indicates that my account of crucial experiments will be formulated within a general account of human judgment. In developing an account of human judgment, one faces a number of alternatives. One is between an account of judgments by individuals or by groups. I shall focus on individuals.[11] Another alternative is whether the individuals in question are to be regarded as 'rational agents' or simply scientists. I shall focus on scientists who are idealized only in the sense that the objects of any theory (e.g., classical mechanics) are idealized, not in the sense of providing normative standards.[12]

On anyone's account, a crucial experiment is designed to decide between two well-defined alternatives. The alternatives may be highly specific hypotheses or more broadly conceived 'approaches' to the same subject matter. The restriction to two alternatives is not as severe as it might seem. Although, in principle, there are always infinitely many logically possible alternatives, in practice scientists rarely face more than a few. And if there happen to be more than two, they can be dealt with in sequence, two at a time.[13]

So the model of crucial decisions to be employed here is a model of an individual scientist trying to decide between two alternative models, which for the moment we will designate simply as M_1 and M_2. This yields the standard two-by-two decision matrix shown in figure 9.3. We need only be a little careful how we understand the alternatives. The label 'M_1' stands for what decision theorists call 'a possible state of the world' and should be understood as referring to the possibility that the world is more or less like the idealized model referred to as 'M_1,' or at least that the world is more like M_1 than M_2. And conversely for M_2. The label 'Choose M_1' means that the agent chooses to regard M_1 as

	M_1	M_2
CHOOSE M_1		
CHOOSE M_2		

9.3 Decision matrix for a choice between alternative models of the same real system.

providing a satisfactory representation of the world, or at least a better representation than that provided by M_2. And conversely for M_2.

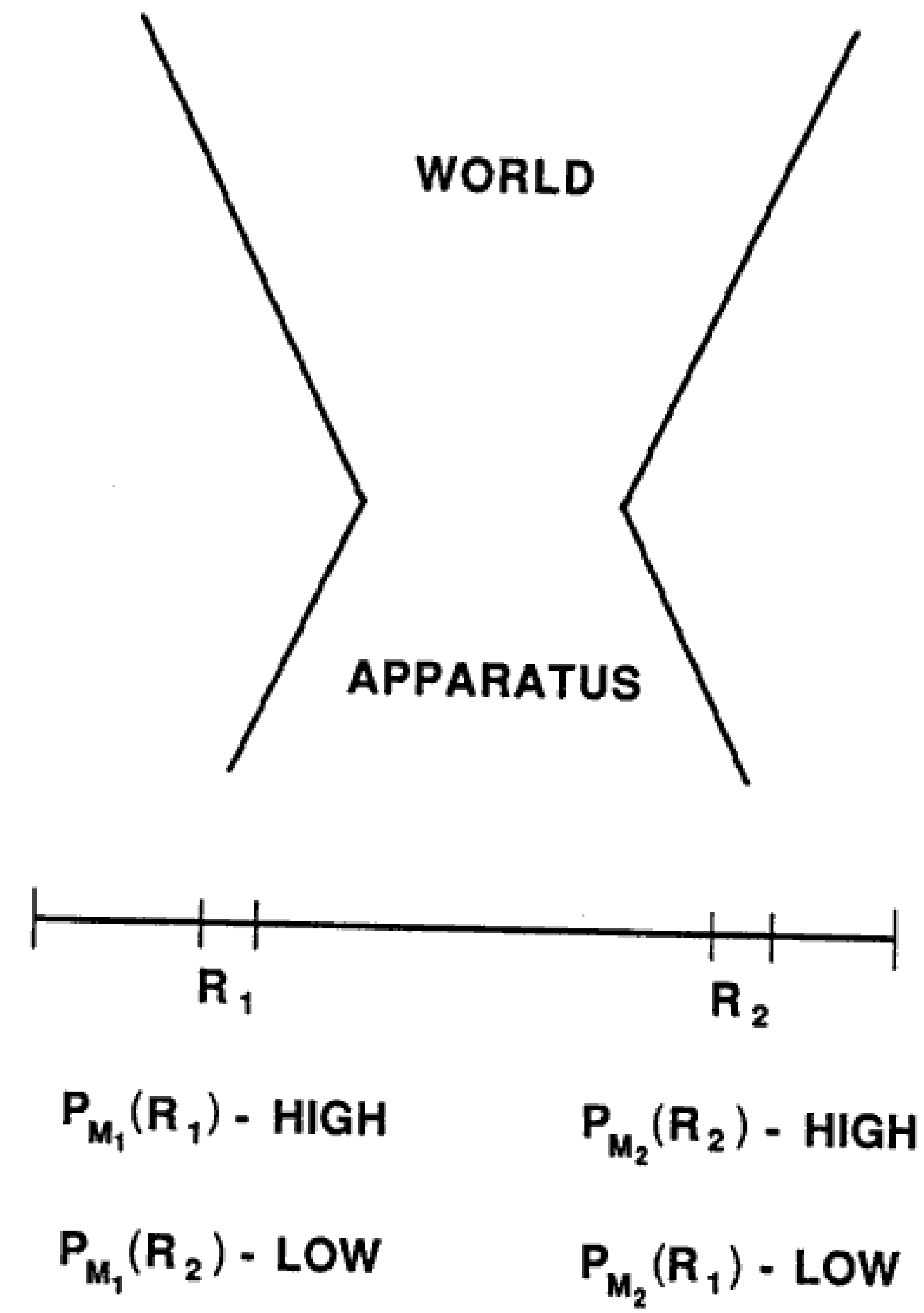
For present purposes we can take an experiment to be a physical process that yields a reading within a specified one-dimensional range of possible readings, as shown in figure 9.4. What makes an experiment *crucial* are the following conditions:

(i) If the actual world is like model M_1, then the experiment is very likely to yield a reading in the range R_1 and very unlikely to yield a reading in the range R_2.
(ii) If the actual world is like the model M_2, then the experiment is very likely to yield a reading in the range R_2 and very unlikely to yield a reading in the range R_1.

The connection between these conditions and the decision matrix is made by the following obvious 'decision rule':

(a) If the experiment yields a reading in the range R_1, choose M_1.
(b) If the experiment yields a reading in the range R_2, choose M_2.
(c) If the experiment yields some other reading, reconsider the whole problem.

That this is the appropriate decision rule can be seen simply by running through the possibilities. If the world really is captured by M_1, then, by condition (i), the experiment will most likely yield a result in range R_1, and, following the decision rule (a), one will choose M_1 as providing the better representation of the world. This is clearly the



9.4 A schematic representation of a crucial experiment.

appropriate choice. Similarly, if the world really is captured by M_2, then, by condition (ii), the experiment will most likely yield a result in range R_2, and, following the decision rule (b), one will choose M_2 as providing the better representation of the world. This is again clearly the appropriate choice. Either way, one is very likely to make the 'right' choice. Of course, if the reading is something else, there are lots of possibilities, including that neither M_1 nor M_2 is a very good representation of the world, that the experiment was badly done, etc.

There are many things remaining to be said about this understanding of crucial experiments, and many things said that might be disputed.[14]

For present purposes only one aspect requires further clarification. The expressions 'very likely' and 'very unlikely' in the two conditions stated above must refer to *physical probabilities* (propensities) in the world. They cannot refer to degrees of belief or epistemic judgments. That would lead to a completely different account of crucial experiments.

In a single spin of a fair roulette wheel, for example, it is very unlikely that the result will be double zero. It is not just that people attach a low degree of belief to this outcome. Rather, given the physical construction and operation of a roulette wheel, a double zero is physically unlikely. The reason most people give little credence to a belief in this outcome being realized is that they know it is physically unlikely. This is not to say that access to knowledge of physical probabilities is mysteriously direct. On the contrary. These judgments, like all other judgments about the world, are based on more or less definite models of the world. So judgments about physical probabilities, like all judgments about the physical world, are *model-based*.

We are now, finally, ready to proceed to the main objective of this paper, which is to show how visual presentations of both models and data can be used in crucial decisions about which models best represent the real world.

5. IMAGES AND ARGUMENTS

At this point I will narrow the discussion to the example of twentieth-century geology.[15] Here the alternative scientific theories are better thought of as broadly conceived *approaches* to geophysics. One approach, commonly labelled *stabilism*, is that the major geological features of the earth, particularly oceans and continents, originally formed in roughly their current configuration and have remained stable in those positions throughout geological time. The overall mechanism was taken to be cooling, contraction, and solidification of an originally molten sphere. The alternative approach, *mobilism*, is that the relative positions of the continents and oceans have altered in major ways in geological time, that is, since the original formation of solid land masses. It is a standard part of mobilism, for example, that the Atlantic Ocean is a relatively recent product of a separation of North and South America from Europe and Africa respectively.

During the 1920s, the mobilist cause was championed by Alfred Wegener, a German scientist whose earlier work was in meteorology and atmospheric physics. Wegener provided mobilism with many

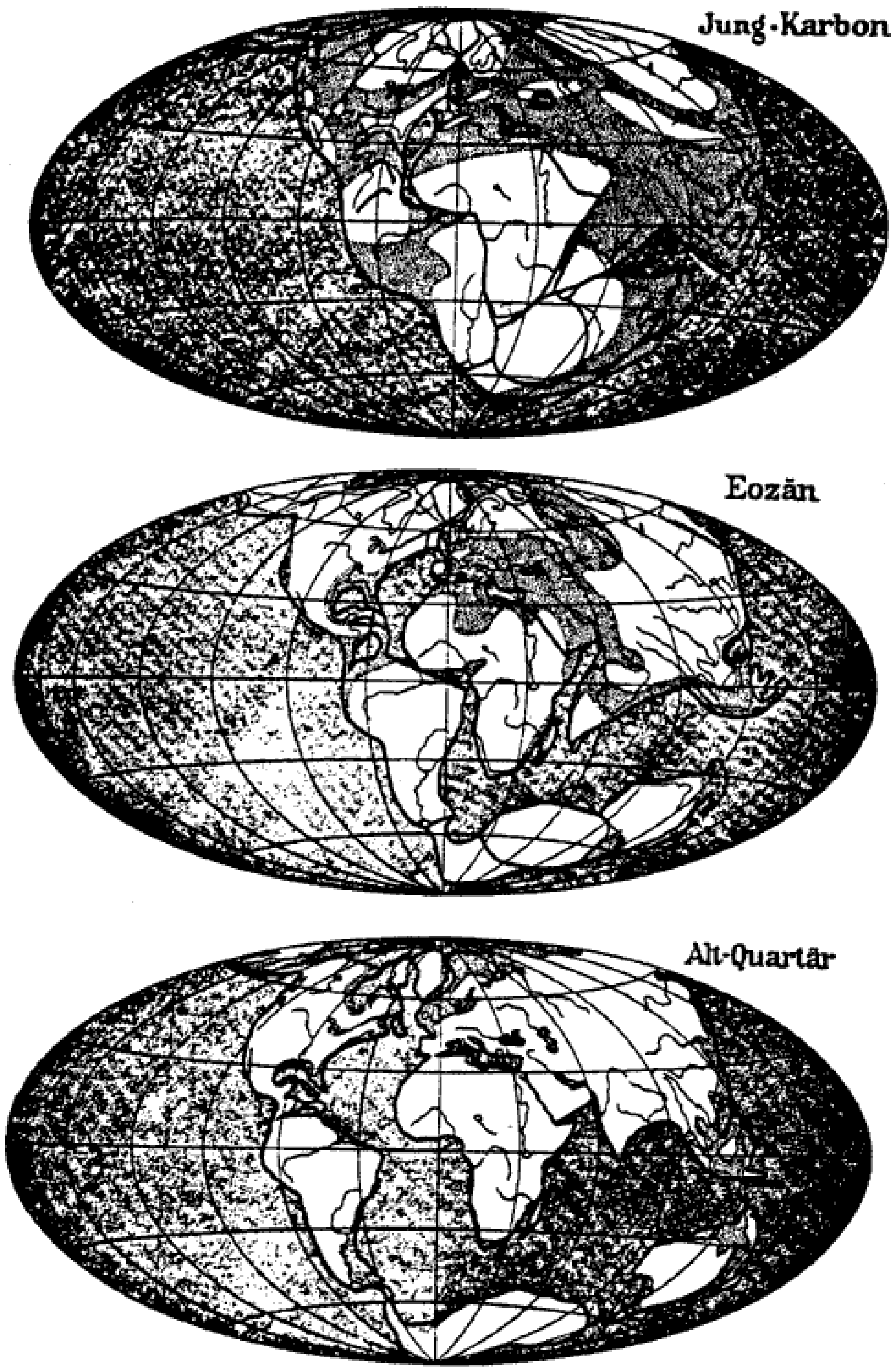
dramatic visual presentations, most notably a series of three world maps picturing the breakup of Gondwanaland, Wegener's original supercontinent containing most of the world's land mass (fig. 9.5). These pictures, which first appeared in the third (1922) German edition of his book, *Die Entstehung der Kontinente und Ozeane*, show the breakup taking place between the Carboniferous (300 million years ago) and the Early Quaternary (500 thousand years ago). I do not claim that these maps constitute the entire content of Wegener's mobilism. Rather, they are visual models which are part of a diverse family of models which all together constitute Wegener's theoretical resources for presenting a mobilist history of the earth.

Wegener gathered evidence for mobilism from many domains, including geology, geophysics, palaeontology, palaeobotany, and palaeoclimatology. Here I will concentrate on just one piece of evidence, the celebrated 'fit' between the eastern coastlines of North and South America and the western coastlines of Europe and Africa. Figure 9.6 reproduces Wegener's sketch of this fit as it appeared in the first (1915) edition of his book. It is a crude sketch. There exist far better drawings exhibiting a better fit dating from over half a century earlier.[16]

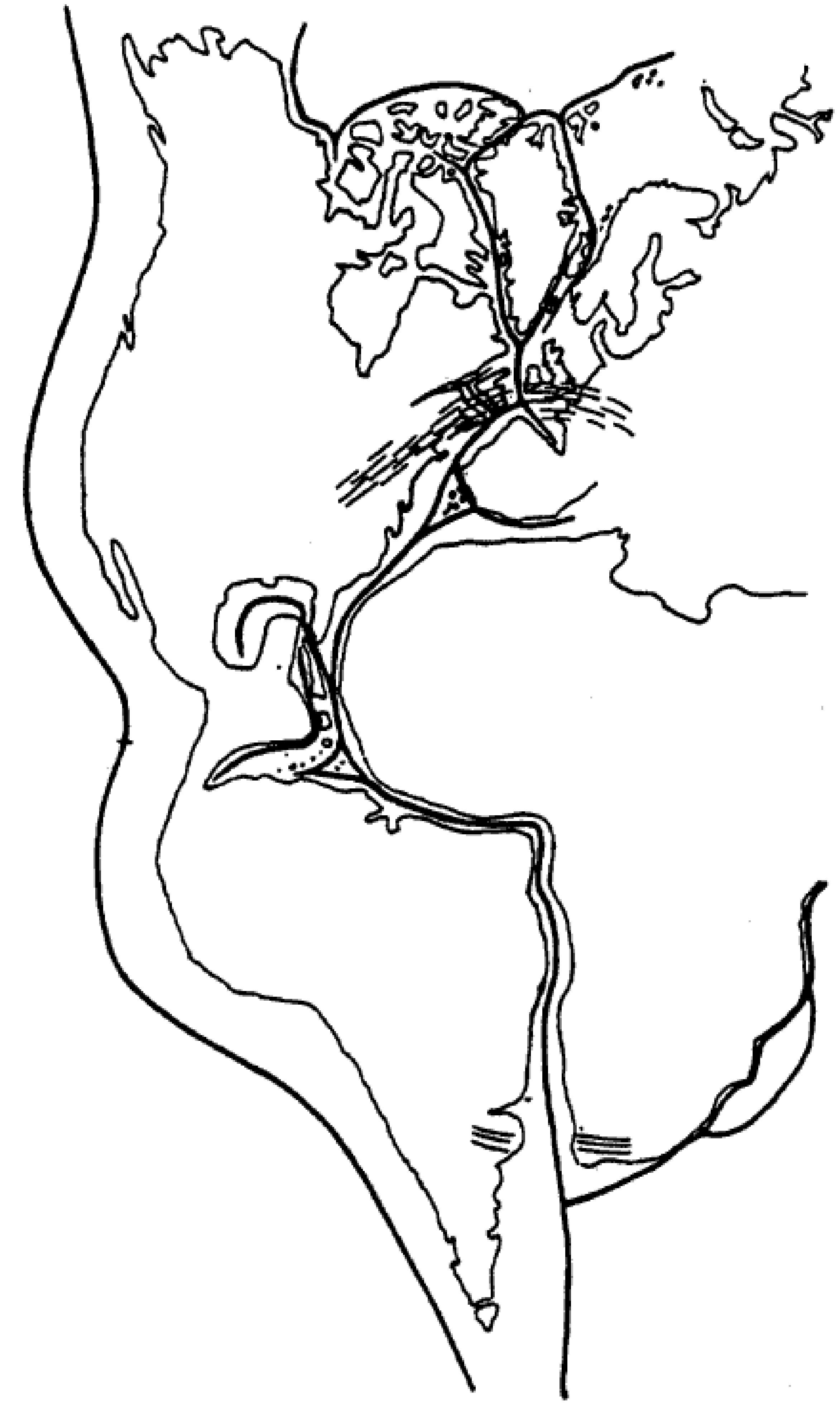
What is notable in this sketch is the explicit attention paid to geological features *other* than the fit of the coastlines. In particular, Wegener has marked areas crossing roughly between England and New England, and between South Africa and southern South America, where mountain ranges appear roughly continuous across the postulated border. These congruences play a major role in his presentation, and apparently did so in his own thinking as well.[17]

Referring to the match in coastlines, Wegener at one point remarks that it reminds him 'of the use of a visiting card torn into two for future recognition' (1924, p. 44). This is a highly visual metaphor. A little later he modifies and expands the metaphor. Referring to both the match in coastlines and the match in features across the boundary, he writes:

> It is just as if we put together the pieces of a torn newspaper by their ragged edges, and then ascertained if the lines of print ran evenly across. If they do, obviously there is no course but to conclude that the pieces were once actually attached in this way. If but a single line rendered a control possible, we should have already shown the great possibility of the correctness of our combination. But if we have *n* rows, then this probability is raised to the *n*th power. (1924, p. 56)



9.5 Wegener's visual representation of the breakup of Gondwanaland (A. Wegener 1922).



9.6 Wegener's sketch of the Atlantic coastlines indicating continuous mountain ranges across the assumed line of separation (A. Wegener 1915).

Here Wegener appears to go beyond my analysis of crucial decisions, claiming a high probability for his theory itself. But his analysis of the evidence includes the two conditions required for my analysis to be operative. That is, if his mobilist account is correct, then the existence of congruences like those noted is highly probable. Conversely, if stabilism is correct, and the continents formed independently of one another, then such congruences are highly unlikely. Since Wegener clearly believes that the congruences do exist, my analysis could explain why he thinks that mobilism is the obvious choice.

In presenting Wegener's argument, I have employed images he himself utilized. What exactly is the role of the images in the presentation? One cannot argue, I think, that the images are *logically* essential. Any information that can be presented in a two-dimensional image can also be presented in a linear, symbolic form, as the digital encoding of images makes obvious. But we are not here concerned with how logically possible scientists might reason. We are concerned with how actual scientists do reason.[18] Wegener's presentation makes it clear both that the images played a large role in his own thinking and that he expected them to play a role in the thinking of his audience as well. But what is that role?

The images, I suggest, function as partial visual models of the relevant features of the earth. As such they provide grounds for model-based judgments about the physical probabilities that would be operative in the world if it were structured according to the model. Thus, the images provide a basis for the model-based judgments regarding physical probabilities needed in my account of crucial decisions.

Of course not all the information necessary for making the required probability judgments is present in the image itself. In Wegener's case, for example, one must know that mountain ranges are relatively rare. They do not exist all up and down the coasts of Europe, Africa, and the Americas. Moreover, mountain ranges have distinctive characteristics. So finding several mountain ranges that are congruent across the boundary when the coastlines are lined up according to their matching shorelines is indeed physically unlikely if those mountain ranges had been formed independently on continents separated by thousands of miles.

What happens in such cases, I suggest, is that the visual model serves as an organizing template for whatever other potentially relevant information the agent may possess, regardless of how that information is encoded. The visual image guides the agent's recall of stored

information by providing a guide to what, within the agent's diverse store of information, is most relevant to the required probability judgments.

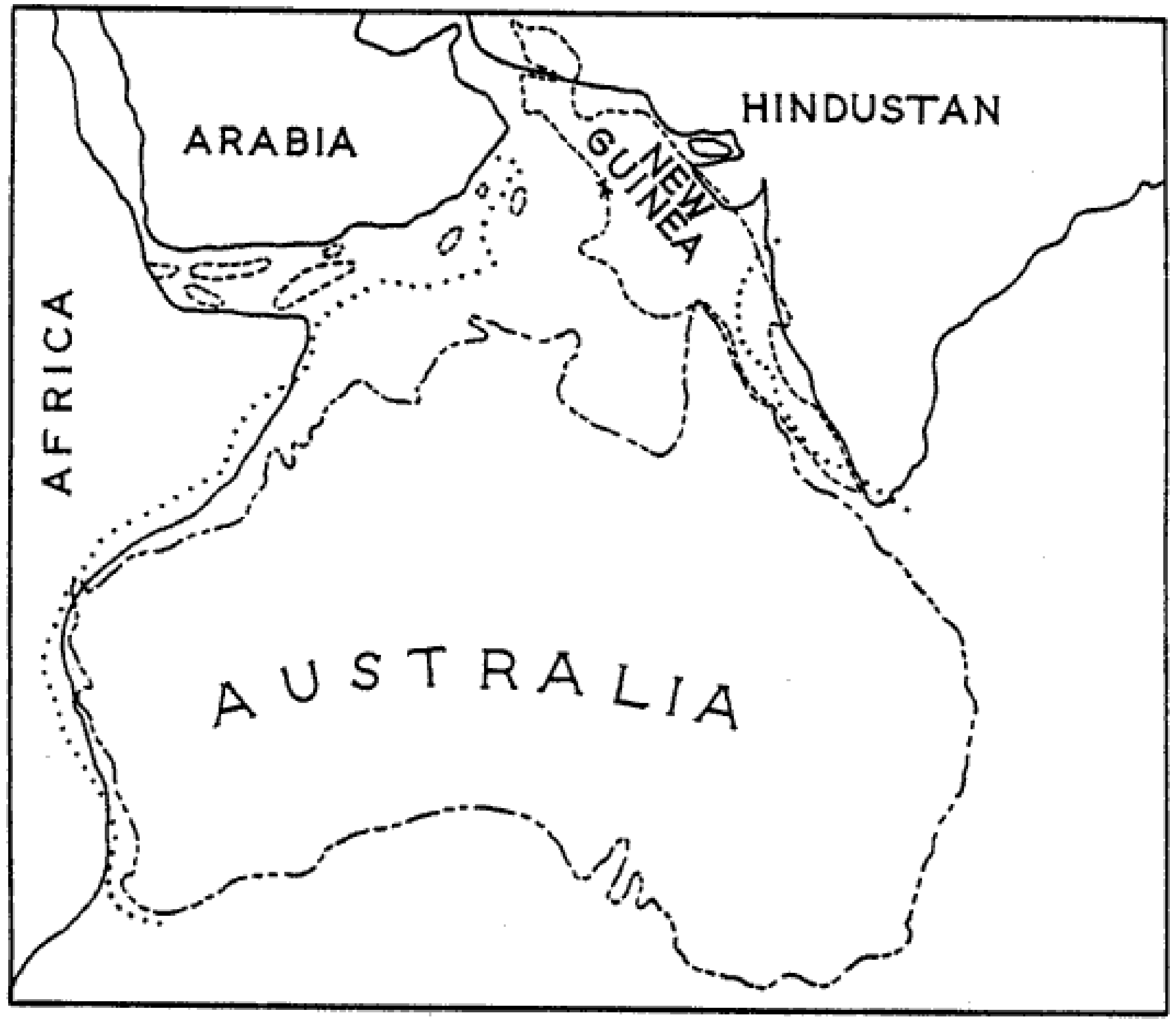
For the purposes of the overall argument of this paper, it is not necessary that my suggestions regarding the role of images be scientifically correct. It would be nice, of course, if something like this were indeed the case. And it is in line with some current thinking in the cognitive sciences. But all my argument requires is that some such account be physically (and psychologically) possible. That shows at least that images could play a significant role in scientific reasoning. And that is enough to refute in principle claims that no such role is possible.

Before leaving this example, I would like to illustrate my position with one further image that played a role in the debates over mobilism in the 1920s. Two years after the 1924 publication of the English translation of Wegener's book, *The Origin of Continents and Oceans*, the American Association of Petroleum Geologists sponsored a symposium in New York City on mobilism (van der Gracht 1928). The symposium featured leading scientists from around the world, including Wegener himself. Of the fourteen participants, roughly one-third supported mobilism, one-third were genuinely open-minded, and one-third were strongly opposed to mobilism.

Among the arguments against Wegener in particular was one offered by Yale geologist Chester Longwell based on the map shown in figure 9.7. This map shows a fairly good fit of the coastlines of Australia and New Guinea within that of the Arabian Sea. But no one present, including Wegener, wished to argue that Australia once filled the Arabian sea.

On my analysis of crucial decisions, the real point of this image is to undermine one of the conditions for Wegener's view of the decision in favour of mobilism. Wegener's position requires that it be physically improbable that there be such matching coastlines within a stabilist model of the earth. Longwell's map provides a clear visual presentation of just such a match. One need not invoke much additional information to be led towards the conclusion that, even within a stabilist model, such matches may not be so improbable as Wegener's position requires. As Longwell himself put it:

> This case is worth some study, in connection with the better known case of South America and Africa, in order to convince ourselves that apparent coincidence of widely separated coast lines is probably accidental wherever



9.7 Longwell's map showing Australia fitting into the Arabian Sea (W. van der Gracht et al. 1928).

found and should not influence anyone unduly in considering the displacement hypothesis. (van der Gracht 1928, p. 153)

In short, there may be visual force on both sides of an argument.

6. THE VISUAL DEVELOPMENT OF THEORETICAL MODELS

Wegener died tragically in 1930 on an expedition to Greenland in search of new evidence for mobilism. About the same time, an English geologist, Arthur Holmes, suggested a new mechanism for mobilism. Inspired by the discovery of natural radioactivity, Holmes reasoned that such radioactivity in the earth might be able to produce sufficient heat to create convection currents of molten minerals just below the earth's

crust. These currents could split the crust and move it laterally great distances before turning downward towards the core. Figure 9.8 reproduces Holmes's visual rendition of this model, in which a continental block is ripped in two creating a new ocean where once land had been.[19]

In spite of its visual power, Holmes's model seems to have done little to stave off the general decline of interest in mobilism following Wegener's death. A good part of the explanation for its lack of immediate influence, I would argue, is that this model provides no basis for a crucial decision between mobilism and stabilism. The processes pictured in Holmes's model would be taking place well below the earth's crust, too remote for any then known instruments. And contemporary surface manifestations, if any, would take place too slowly to measure. Holmes himself did not even conjecture a possible crucial experiment.

Holmes's convection model was revived thirty years later by the American geologist Harry Hess. Figure 9.9 reproduces Hess's version of the model (Hess 1962, p. 607). The main difference is that Hess has the convection current rising under the ocean floor rather than a continental block. This was because Hess, unlike Holmes, intended his model to explain the origin of the great ocean ridge systems first explored in the 1950s. The ridges, on Hess's model, are produced directly above the rising convection current, which then spreads out creating a new sea floor. But Hess's model, like Holmes's, provides no basis for a crucial decision between mobilism and stabilism.

The makings of a crucial experiment were provided by a new graduate student in geophysics at Cambridge, Fred Vine, and his recently appointed supervisor, Drummond Matthews. In late 1962, Matthews returned from an expedition to the Indian Ocean, where he had obtained systematic measurements of the total magnetic field at the level of the ocean floor across the Carlsberg Ridge. The task of analysing this magnetic data fell to Vine, while Matthews went off on his honeymoon.

Using then very new computer techniques, Vine determined that the magnetic readings across the ridge showed a small periodic variation as one moved away from the centre of the ridge. Similar periodic variations in magnetic intensity along the ocean floor had earlier been observed in other areas, such as the Pacific Ocean off the coast of North America. Taking these variations in magnetic intensity as a real phenomenon requiring explanation, Vine, early in 1963, set about finding one.

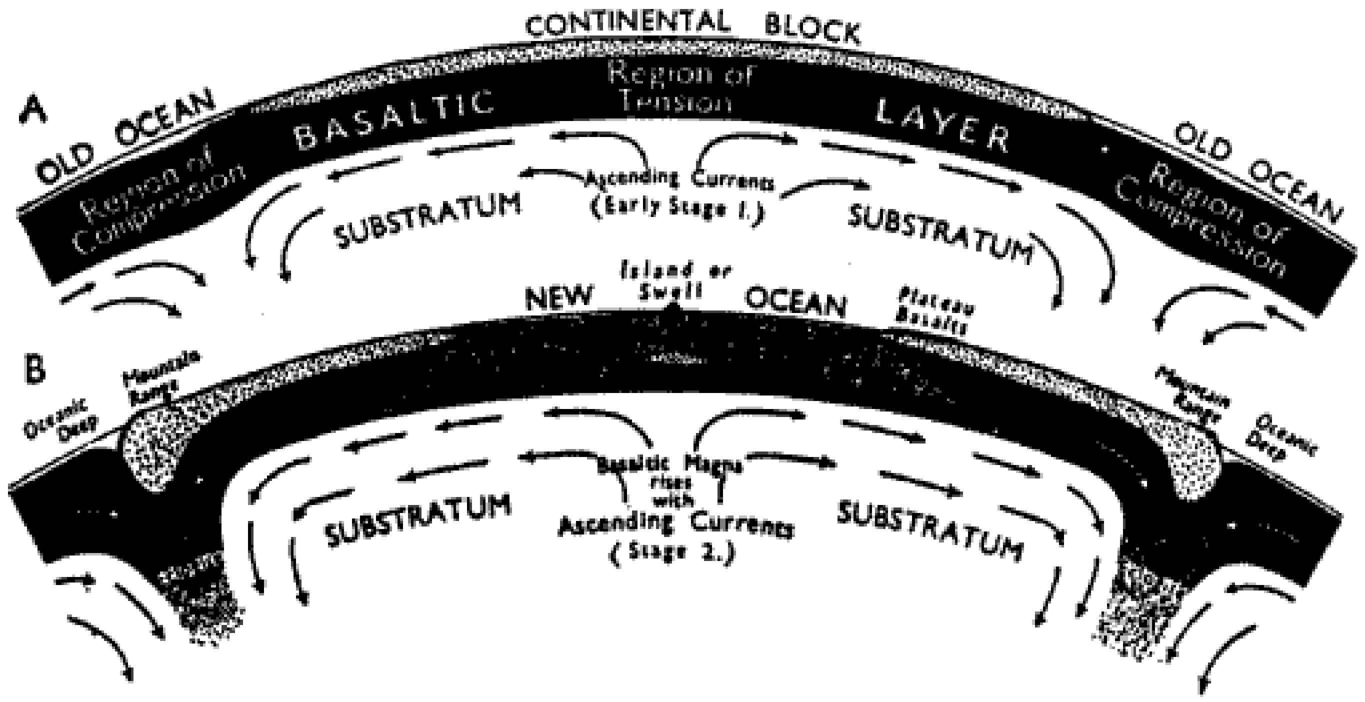
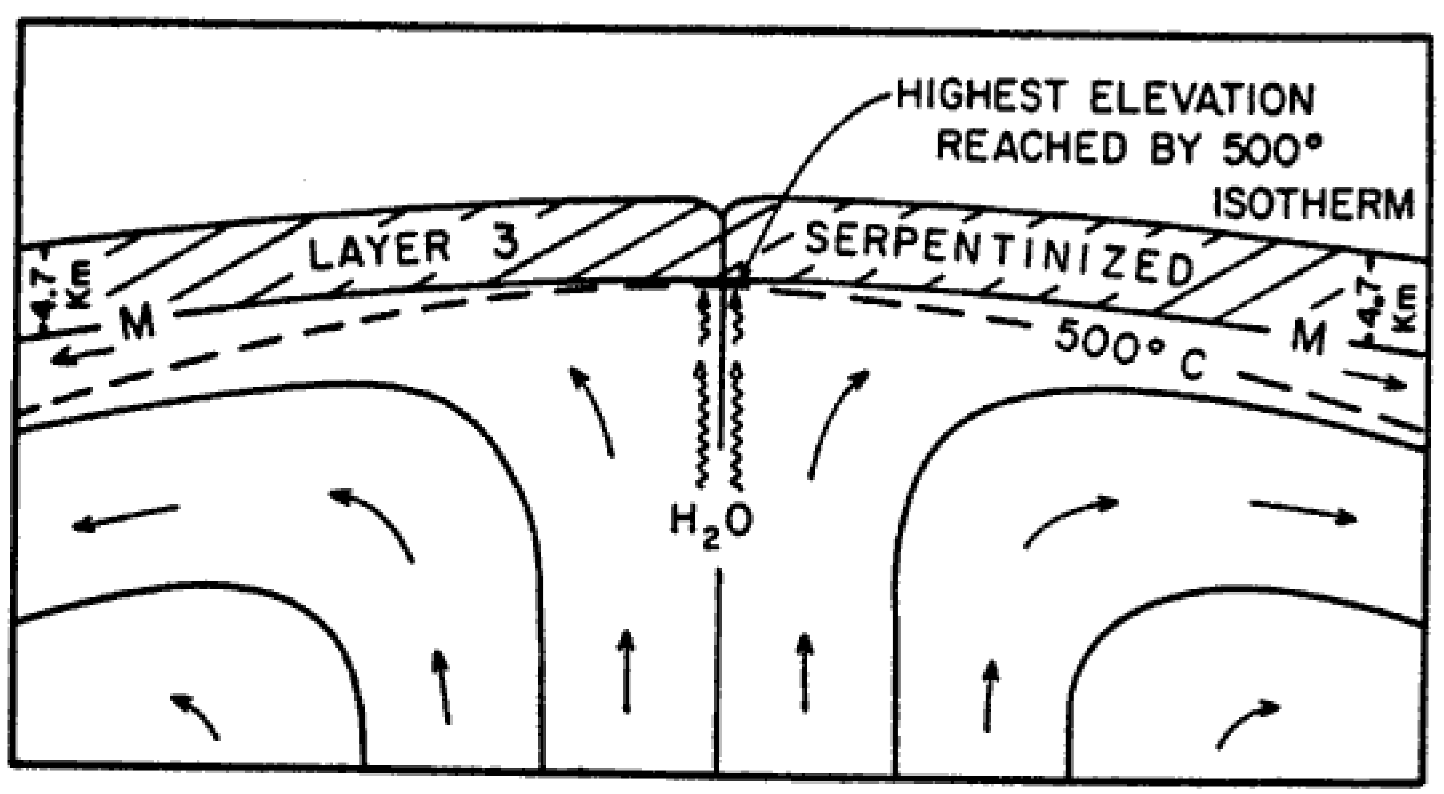


FIG. 262

Diagrams to illustrate a purely hypothetical mechanism for "engineering" continental drift. In A sub-crustal currents are in the early part of the convection cycle (Stage 1 of Fig. 215). In B the currents have become sufficiently vigorous (Stage 2 of Fig. 215) to drag the two halves of the original continent apart, with consequent mountain building in front where the currents are descending, and ocean floor development on the site of the gap, where the currents are ascending

9.8 Holmes's dynamic visual representation of convection currents splitting a continent to produce a new ocean (A. Holmes 1944).



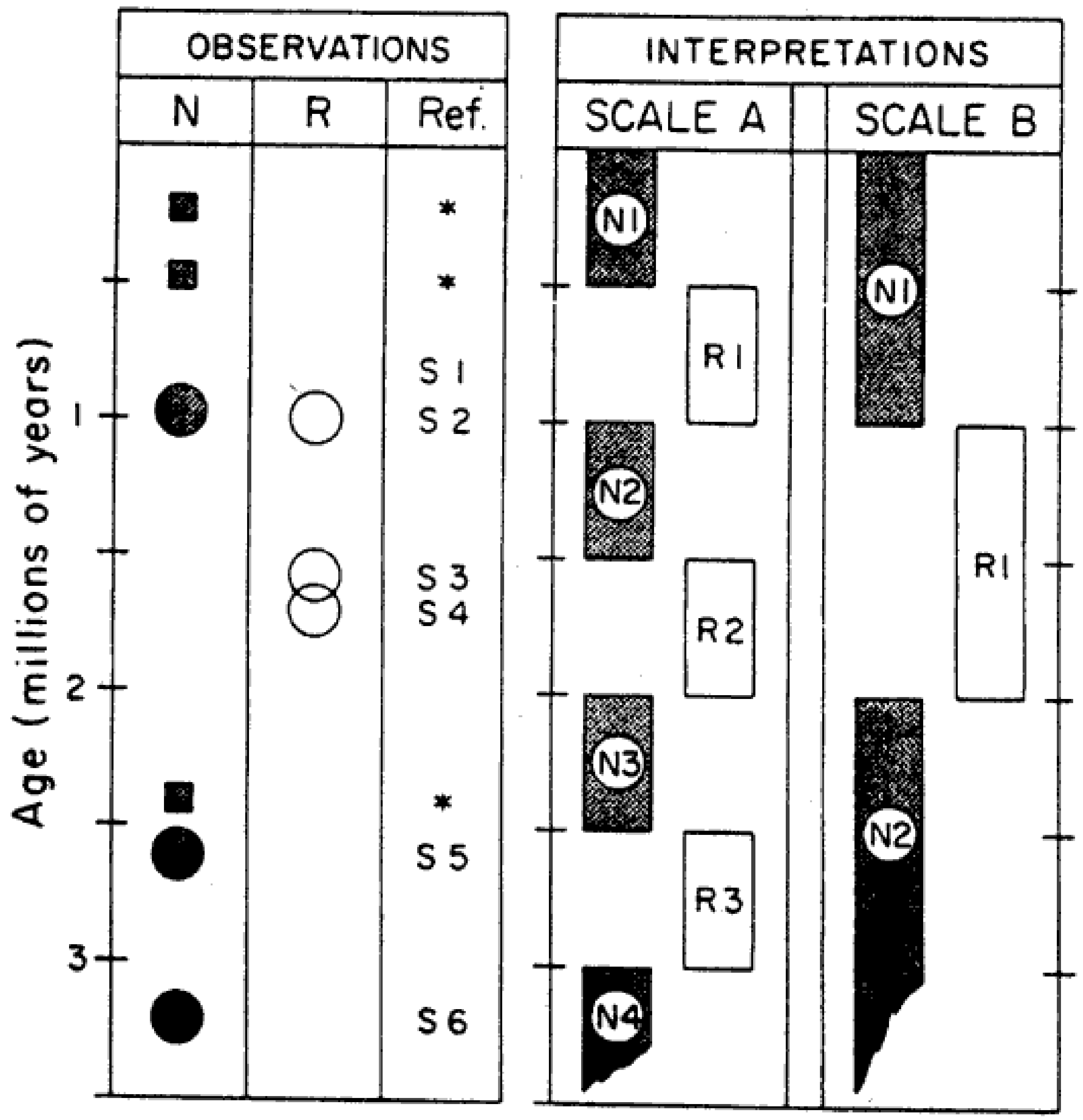
9.9 Hess's dynamic visual model of sea-floor spreading produced by convection currents (H.H. Hess 1962).

Vine was keenly aware of Hess's model of sea-floor spreading, having seen Hess himself present it during a conference at Cambridge in January 1962. A possible link to the observed variations in magnetic intensity was provided by initially unrelated work on palaeomagnetism. Researchers in California, led by Allan Cox, had been examining the direction of remanent magnetism in core samples from lava flows. Such samples provide a measure of the direction of the earth's magnetic field when the examined material was molten since magnetic material in a molten fluid would tend to line up with the existing magnetic field of the earth. What they found, confirming scattered findings dating to half a century earlier, was an apparent change in direction of the earth's magnetic field several times in the past four million years.

Figure 9.10 shows a visual presentation of both the data and the theory in one of the first publications, in mid-1963, of the California group (Cox, Doell, and Dalrymple 1963). The single vertical scale, representing age in millions of years, starts with zero at the top and shows increasing time into the past as one moves down the scale. This arrangement represents the obvious geological fact that in a lava flow the younger materials from recent eruptions are towards the top while the older materials are deeper. Each data point represents a number of rock samples from a given site, with the average age of the samples indicated by the location of the data point relative to the vertical scale. The polarity of the sample, 'normal' or 'reversed,' is indicated by its location in the left or right column. The rival models are simple. They just represent the magnetic field of the earth as having been continuously normal for a time into the past, then being reversed, then being normal, and so on in equal time intervals. One model puts the period of the reversals at a half million years, the other at a million. That both models are consistent with the data can be seen immediately in the graphical presentation.

What is the connection between (i) Vine's data showing regular variation in magnetic field intensity extending out from an ocean ridge, (ii) Hess's model of sea-floor spreading, and (iii) evidence for geomagnetic reversals? The answer cries out for a dynamic, visual model, but none seems to have been published during the crucial years 1963–6. Vine and Matthews's 1963 paper contains, instead, the following verbal description:

> The theory is consistent with, in fact virtually a corollary of, current ideas on ocean floor spreading and periodic reversals in the Earth's magnetic field. If the main crustal layer ... of the oceanic crust is formed over a convective up-



9.10 The first visual presentation of data and models by the California group investigating geomagnetic reversals (A. Cox et al. 1963).

> current in the mantle at the centre of an oceanic ridge, it will be magnetized in the current direction of the Earth's field ... Thus, if spreading of the ocean floor occurs, blocks of alternately normal and reversely magnetized material would drift away from the centre of the ridge and parallel to the crest of it. (Vine and Matthews 1963, p. 948)

No one can deny, however, that in reading this description it helps to refer back to the dynamic visual models of Holmes and Hess.[20]

Following the above description is a visual presentation of the magnetic data and a corresponding model of the sea floor across three different ridges (fig. 9.11). This appears to be an adaptation of Cox's model for geomagnetic reversals, except that the blocks of alternately magnetized material are laid out horizontally rather than vertically. This difference, of course, reflects the differing causal processes suggested as having produced the two configurations of differentially magnetized materials.

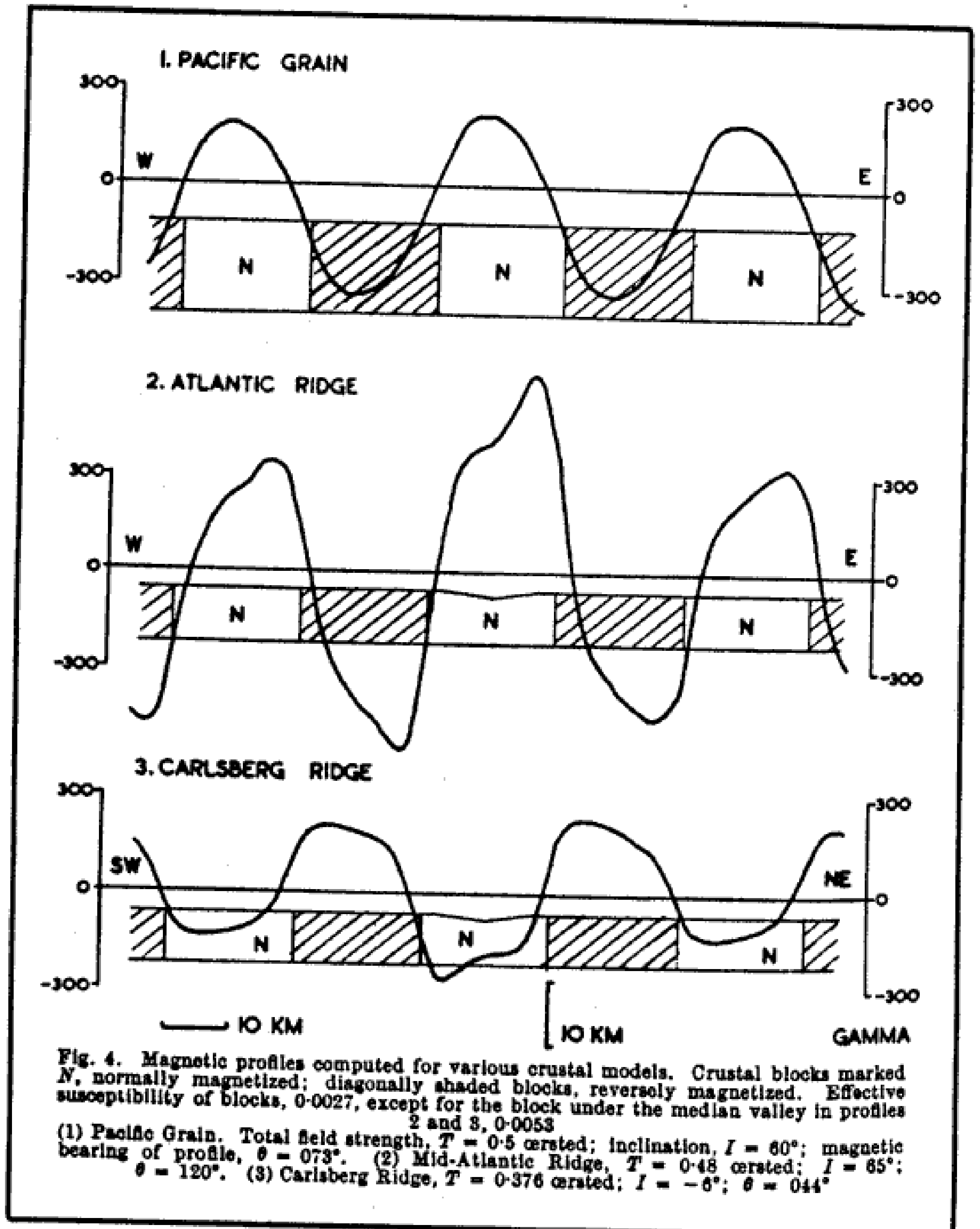
Publication of Vine and Matthews's paper seems to have convinced almost no one of the reality of sea-floor spreading and the mobilism it implies. Not even they were willing to claim they had proven the case. In the last few lines of their 1963 article, they write:

> It is appreciated that magnetic contrasts within the oceanic crust can be explained without postulating reversals of the Earth's magnetic field; for example, the crust might contain blocks of very strongly magnetized material adjacent to blocks of material weakly magnetized in the same direction. However, the model suggested in this article seems to be more plausible because high susceptibility contrasts between adjacent blocks can be explained without recourse to major inhomogeneities of rock type within the main crustal layer or to unusually strongly magnetized rocks. (Vine and Matthews 1963)

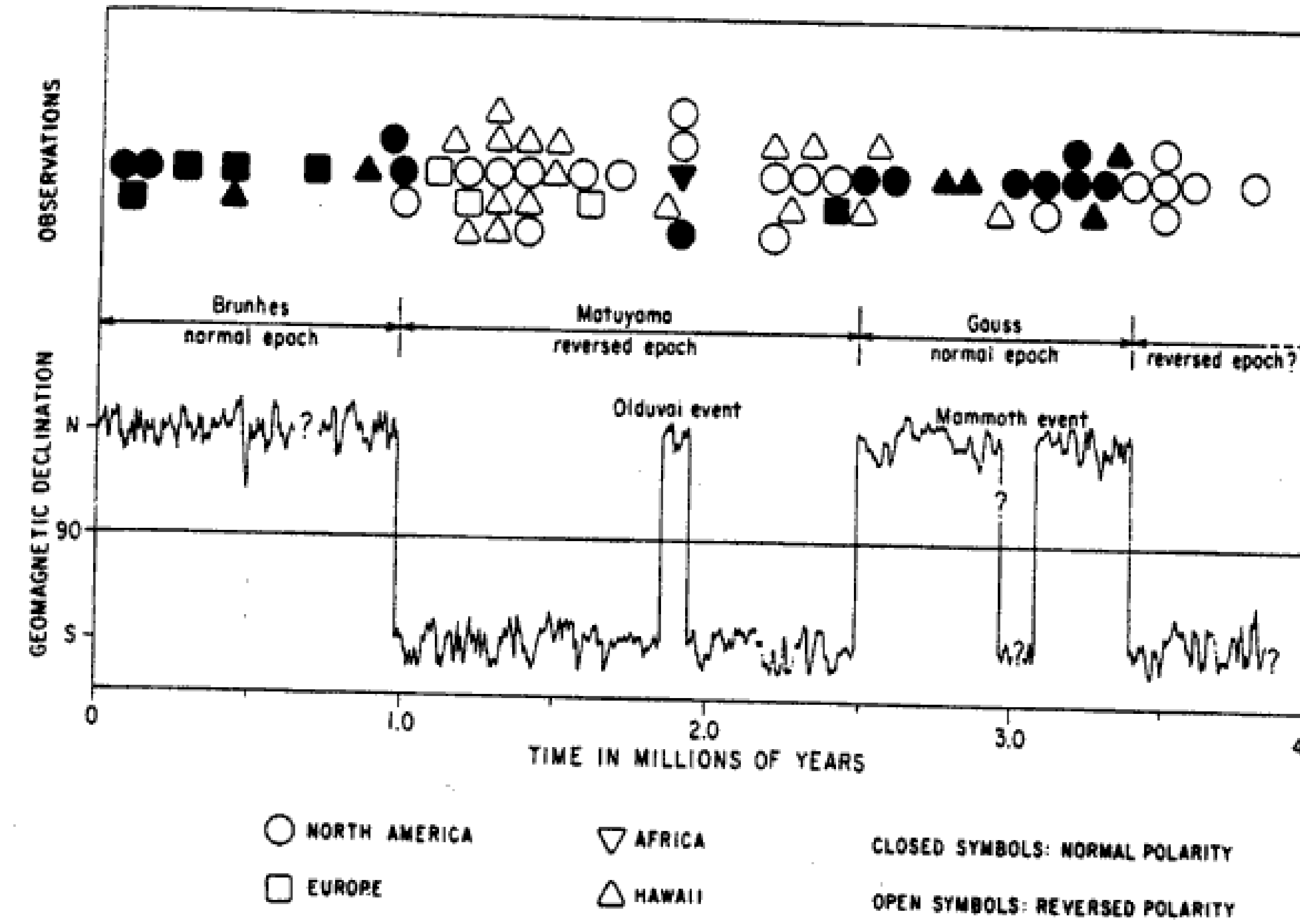
In terms of my model of crucial decisions, they do seem to think that condition (i) is satisfied. The results obtained are fairly probable given a model incorporating sea-floor spreading and geomagnetic reversals. But these results are not wildly improbable if those assumptions are mistaken and stabilism is correct. So there is no adequate basis for making a crucial decision in favour of sea-floor spreading and mobilism.

Of course, the noted possibilities for stabilist explanations of the data are not directly contained in the visual presentation of their model of the sea floor or of their data. But realizing that the simple periodic structure of the model was just read off the similarly simple periodic structure of the data makes it easy visually to assimilate suggested alternative models. So the visual presentation facilitated the judgment that the prospects for a crucial decision were not yet compelling.

During the next two years, the conditions for a crucial decision improved in one respect, but declined in another. In 1964 the California group (Cox, Doell, and Dalrymple 1964), having acquired data from several new sites, published a new scale of geomagnetic reversals (fig. 9.12). Two major differences from the earlier scale are immediately evident. First, they have given up the assumption of equal time periods of normal and reversed polarity. The major normal and reversed 'epochs' are now of irregular duration. Second, they have



9.11 Vine and Matthews's models for magnetic profiles near ocean ridges (F.J. Vine and D.H. Matthews 1963).



9.12 The second visual presentation of data and models by the California group investigating geomagnetic reversals (A. Cox et al. 1964).

refined the scale to include several brief (100 thousand year) 'events.' The Olduvai event, around 1.9 million years ago, is a brief period of normal magnetism within a long epoch of reversed magnetism. Similarly, the Mammoth event, around 3 million years ago, is a brief period of reversed magnetism within an epoch of normal magnetism.

Meanwhile, Vine and a visiting senior Canadian geologist, Tuzo Wilson, were busy analysing magnetic data from yet another ridge system, this one in the Pacific off the coast of Vancouver. Figure 9.13 reproduces one of their visuals as published in 1965 (Vine and Wilson 1965). Following the California group, their models now exhibit both unequal epochs and several briefer events. In terms of my own model for crucial decisions, the good news was that an aperiodic pattern of reversals with intervening small events is very unlikely to appear in scattered places around a stabilist earth. On any stabilist model, it would take a near miracle for the possible sources of magnetic variation

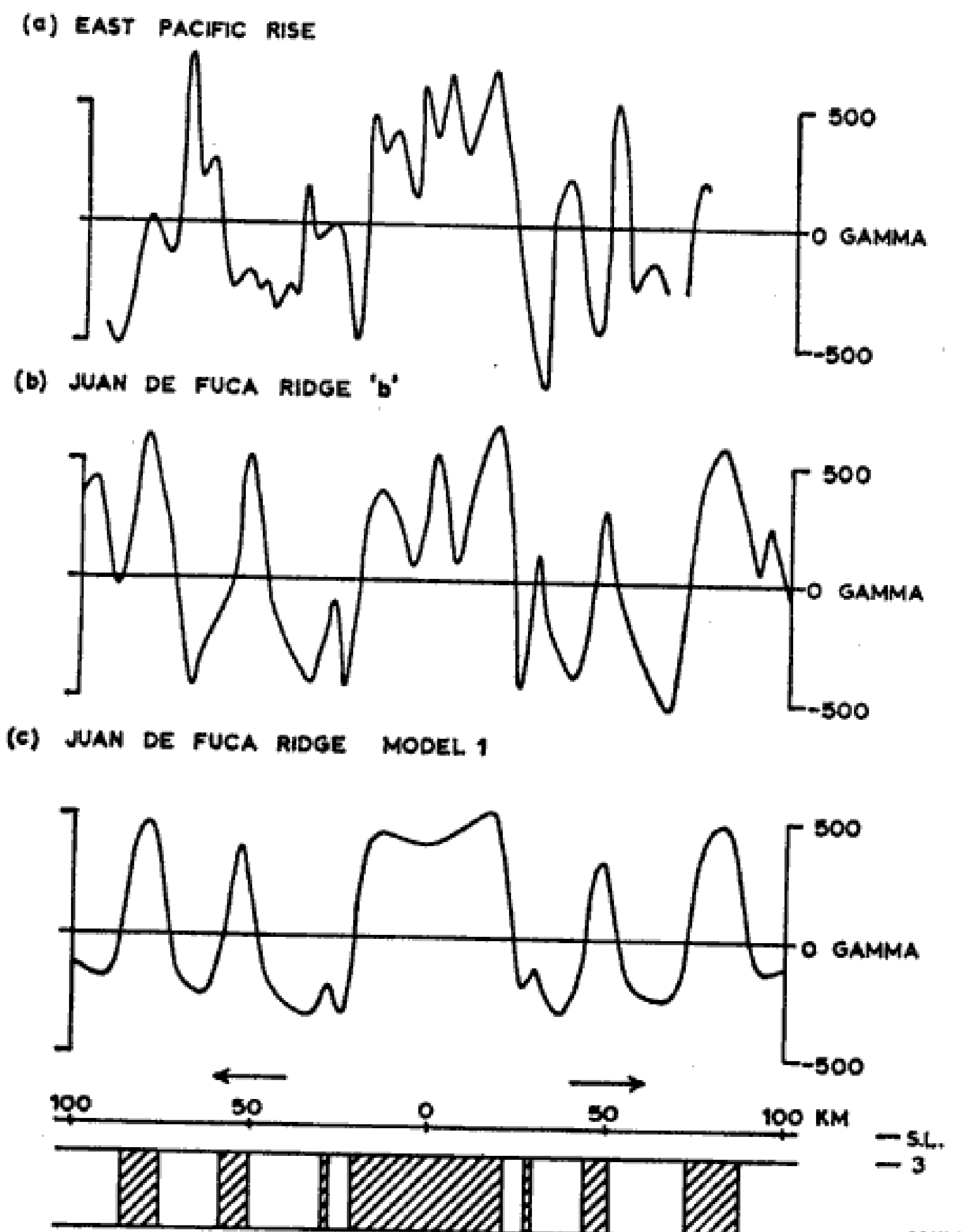


Fig. 2. (*a*) Observed profile across the East Pacific Rise at 59°S, 149°W (*16*). (*b*) Observed profile "b" across the Juan de Fuca Ridge (see Fig. 4). (*c*) Model and calculated anomaly for Juan de Fuca Ridge, assuming generalized crustal blocks (compare Fig. 1*c*).

9.13 Vine and Wilson's visual comparison of magnetic data and crustal model for the Juan de Fuca Ridge (F.J. Vine and J.T. Wilson 1965).

in the sea floor near ridges to produce the same complicated irregular pattern near several different ridges lying in different oceans.

The bad news was that the pattern they were finding in the magnetic sea floor data was not exactly what the Cox scale would lead one to expect if the Vine-Matthews model were correct. If one holds to the constraint that the spreading rate of the sea floor has been roughly constant, there was no way to match up the observed magnetic readings across the ridge with the new scale of reversals published by the California group. The various epochs and events simply did not line up as expected. What was gained in the satisfaction of one condition was lost in failure to satisfy the other.

7. THE PERSUASIVE POWER OF IMAGES

Within a year, the situation had changed dramatically. In late 1965 a research vessel operated by the Lamont Geological Observatory of Columbia University returned from a new geological survey of the Pacific-Antarctic Ridge with the dramatic magnetic profile shown in figure 9.14 (Pitman and Heirtzler 1966, p. 1166). Whereas earlier profiles and geomagnetic time scales had extended out to around four million years ago, this profile extended out a distance corresponding to ten million years, revealing a continuing pattern of reversals never before detected.

The bilateral symmetry of the profile is of particular significance, as is the method used to make it visually obvious. The centre profile shows the magnetic readings moving from west to east at the right of the diagram. The top profile is just the middle profile reversed, with west on the right of the diagram. Merely by scanning visually across the diagram and comparing these two profiles, one can see just how amazingly symmetrical the profile is. That it should be symmetric is an immediate consequence of the Vine-Matthews model, since the sea floor should spread out equally on both sides of a ridge. The lower profile is derived from the model shown at the bottom of the diagram.

About the same time as the data from the Pacific-Antarctic Ridge were being analysed, another group at Lamont was busy analysing the magnetic orientation of sedimentary materials in core samples taken from the ocean floor near the tip of South America. Like cooling lava, sediment traps magnetic materials in their existing spatial orientation as the sediment packs more tightly. If the pattern of geomagnetic reversals is real, it should also be recorded in such sediments. Again the

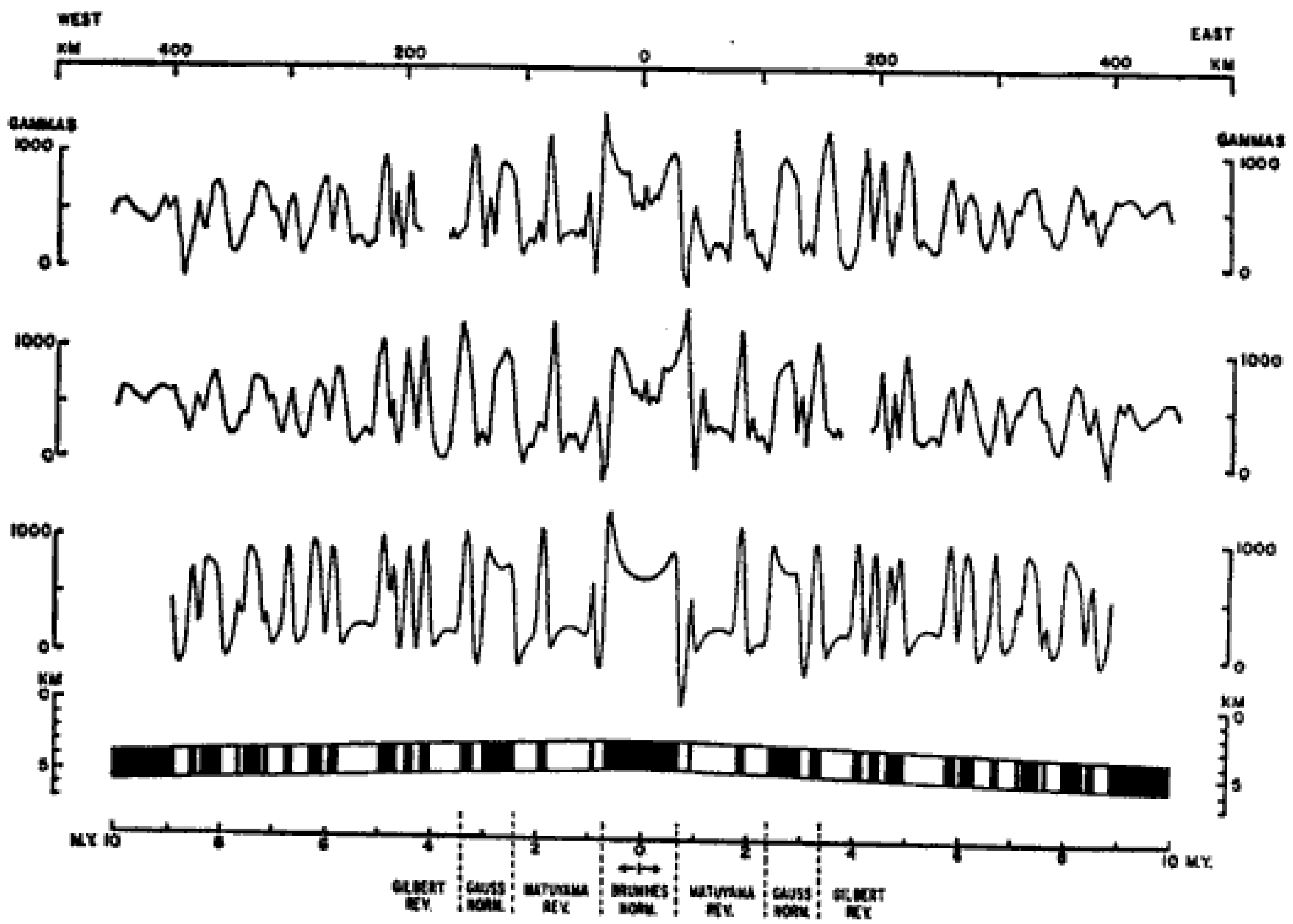


Fig. 3. The middle curve is the *Eltanin*-19 magnetic-anomaly profile; east is to the right. The upper anomaly profile is that of *Eltanin*-19 reversed; west is to the right. On the bottom is the model for the Pacific-Antarctic Ridge. The time scale (millions of years ago) is related to the distance scale by the spreading rate of 4.5 cm/yr. The previously known magnetic epochs since the Gilbert epoch are noted. The shaded areas are normally magnetized material; unshaded areas, reversely magnetized material. Above the model is the computed anomaly profile.

9.14 A visual comparison of magnetic profiles of the Pacific-Antarctic Ridge with a corresponding model. Note especially the symmetry described in the text (W.C. Pitman and J.P. Heirtzler 1966).

evidence, as presented visually in figure 9.15, is dramatic (Opdyke et al. 1966, p. 350). Just by inspecting the diagram, one can see almost immediately that the pattern formed by regions of normal and reversed magnetism within the core samples closely matches that of the magnetic profiles across ocean ridges.

But what of the mismatch between the geomagnetic times scales of the California group and the sea floor profiles which had plagued Vine and Wilson just one year earlier? That too was resolved. Working with rock samples discovered near Jaramillo Creek in New Mexico, several members of the California group discovered that the current period of normal magnetism extended not one million years into the past, but

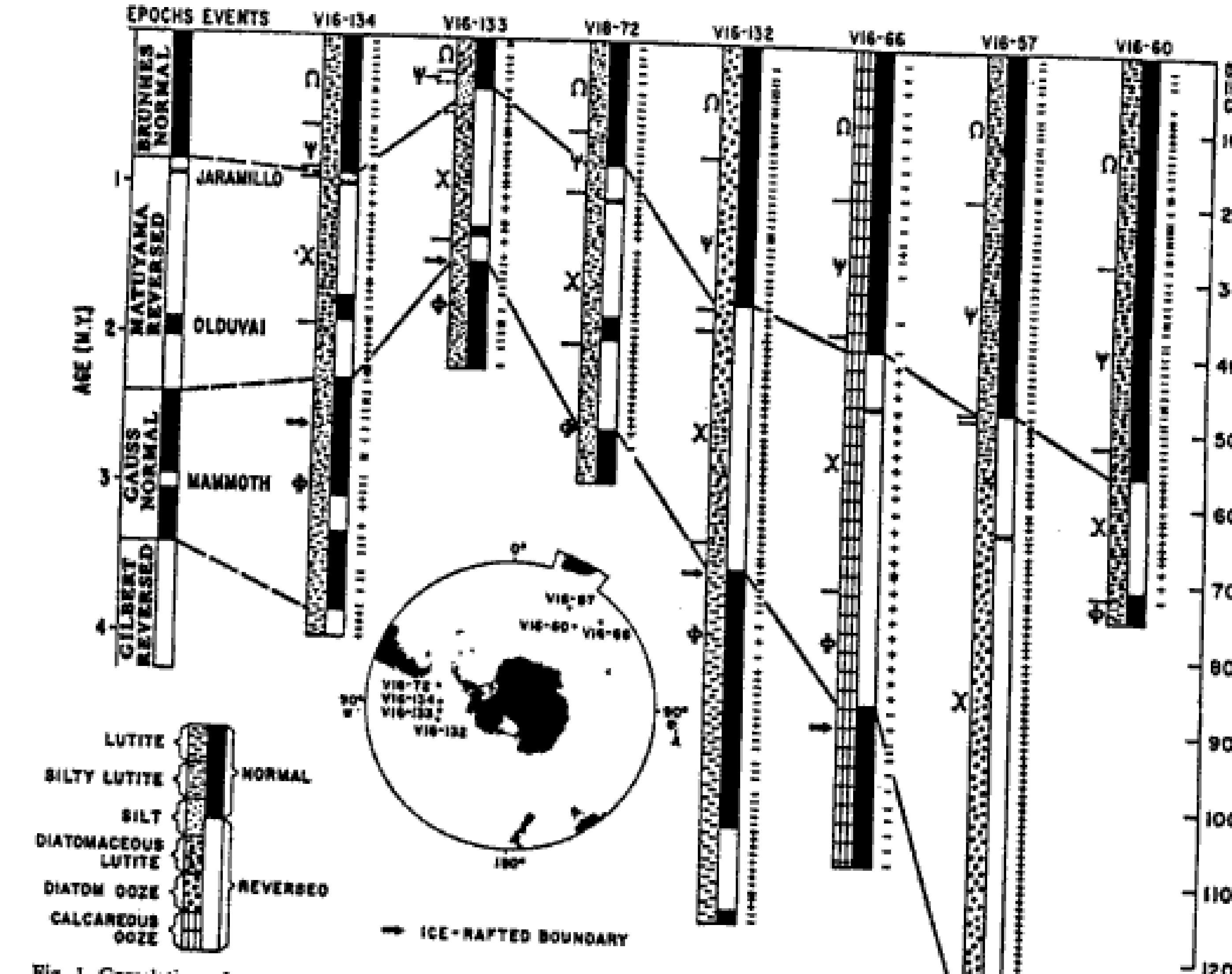


Fig. 1. Correlation of magnetic stratigraphy in seven cores from the Antarctic. Minus signs indicate normally magnetized specimens; plus signs, reversely magnetized. Greek letters denote faunal zones (*17*). Magnetic stratigraphy in left-hand column is from (*1*, *2*). Inset: sources of cores.

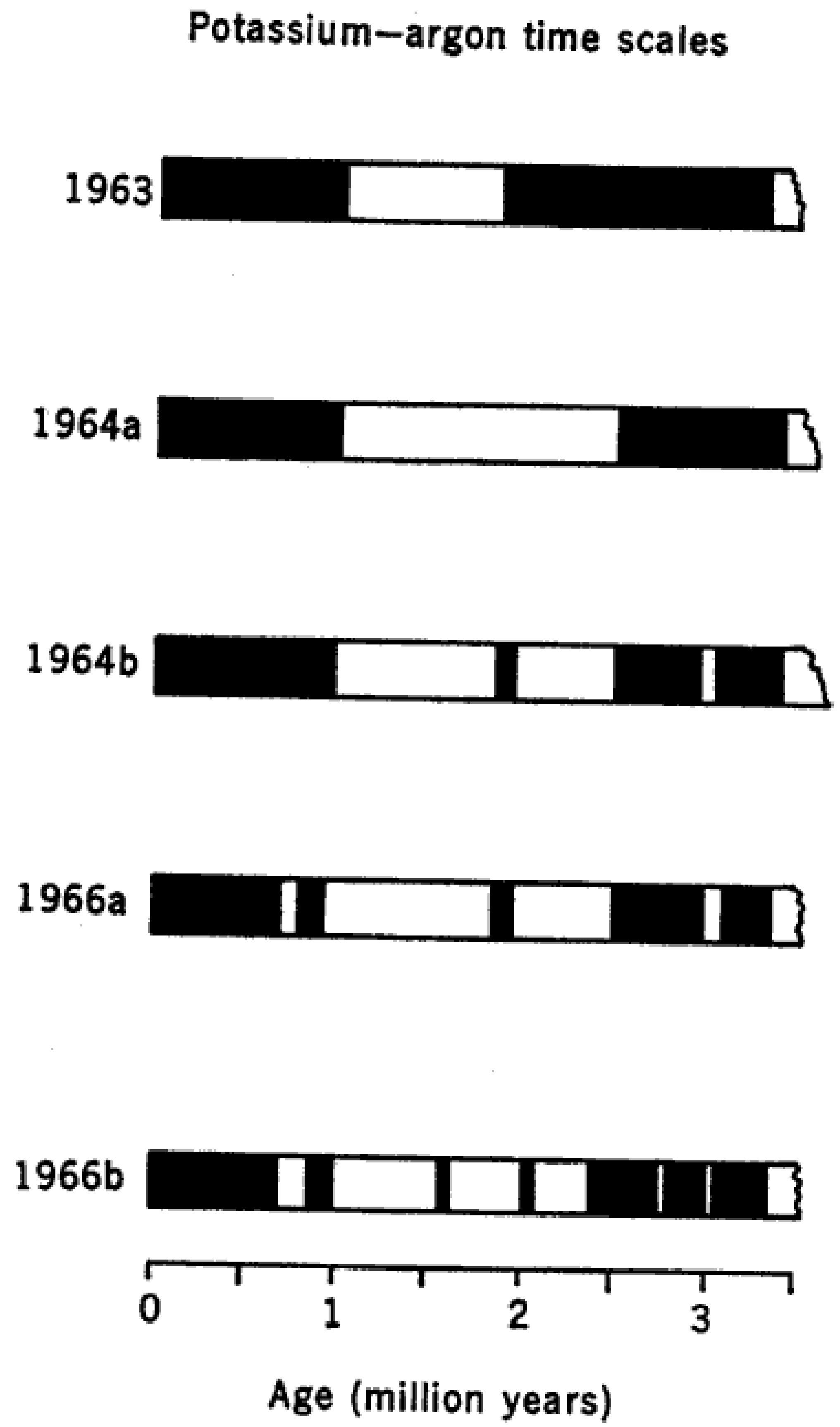
350

SCIENCE, VOL. 154

9.15 A visual comparison of magnetic reversals in deep-sea sediments with the time scale for reversals in terrestrial lava flows (N.D. Opdyke et al. 1966).

only about 0.7 million years. It was followed by a brief period of reversed magnetism and then an event of normal magnetism extending between 0.9 and 1.0 million years ago. Vine and Wilson had gone astray because they had identified the first normal event in their profiles as the nearly two-million-year-old Olduvai event rather than the one-million-year-old Jaramillo event. Figure 9.16 presents Cox's retrospective summary of his group's results during the crucial years 1963–6 (Cox 1969, p. 239).

Versions of these last three images were all presented in a historic session of the April 1966 meeting of the American Geophysical Union. By all accounts, the effect was dramatic. And so it should have been if my account of crucial decisions is correct. That the data obtained were



9.16 Cox's summary of refinements in the California group's scale for geomagnetic reversals during the years 1963–6 (A. Cox 1969).

to be expected if the Vine-Matthews model is correct was well known. What made these presentations especially dramatic was that they showed how utterly improbable the data would be on any stabilist model. What stabilist process (short of divine creation) could possibly have produced that visually dramatic and detailed signature pattern simultaneously in widely scattered continental lava flows, deep sea sediments, and the floor of several different oceans? This was visually obvious to all, regardless of their particular research specialties. As Allan Cox, who chaired the April AGU session, later summed it up, '... there was just no question any more that the seafloor-spreading idea was right' (Glen 1982, p. 339).

8. CONCLUSION

This paper connects several recent themes in the philosophy of science, and in science studies more generally: (i) a model-based picture of scientific theories; (ii) a naturalistic account of crucial decisions; and (iii) interest in the use of visual images in scientific thinking. To structure the paper I assumed a model-based account of scientific theories and used the fact that it could accommodate visual information to support my naturalistic account of crucial decisions. But the argument need not have been structured this way. Probably the most appropriate conclusion is that these three themes are mutually reinforcing, and together support a move away from an exclusive reliance on propositional modes of analysis for understanding the workings of modern science.

NOTES

1 The original version of this paper, under the title 'The Visual Presentation of Theory and Data: A Cognitive View,' was presented at a meeting of the Society for Social Studies of Science in November 1987, and for the Committee on History and Philosophy of Science at Johns Hopkins in December. In February 1989 a later version, under its present title, was presented at the Science Studies Units of the Universities of Bath and Edinburgh, and for the Department of Logic, Methodology and Philosophy of Science at the University of London. A still later version, 'Visual Models in Science: Lessons from the Revolution in Geology,' was presented for the Centenary Conference in the History of Science at the University of Oklahoma in September 1990. The current version, once

again extensively rewritten, owes much to colleagues at all these institutions. I am especially thankful to Professor Brian Baigrie for providing the opportunity for me finally to put these ideas into print. The support of the National Science Foundation is also gratefully acknowledged.

2 Here it is worth noting that Kuhn's book contains not a single illustration.

3 For good examples of the initial philosophical reaction to Kuhn's work, see Shapere 1964 and Scheffler 1967. Kuhn's implicit reply is scattered throughout the essays reprinted in *The Essential Tension* (1977).

4 The now classic references to the new sociology of science include Barnes 1974, Bloor 1976, Latour and Woolgar 1979, Knorr-Cetina 1981, Collins and Pinch 1982, and Collins 1985. This early phase is nicely summarized, with many references, in Barnes and Edge 1982 and by Shapin 1982. For more recent developments, see Latour 1987, 1989, Woolgar 1988, and the contributions in Pickering 1992.

5 For these developments within the sociology of science, see Lynch and Woolgar 1990, particularly the editors' introduction and the essays by Latour and Tibbetts. I myself have reviewed these essays in some detail (see Giere 1994a).

6 The roots of the model-based view go back to the work of J.C.C. McKinsey, Evert Beth, and John von Neumann in the 1930s, '40s, and '50s. It came to prominence in the philosophy of science in the 1960s, '70s, and '80s through their followers, Patrick Suppes, Bas van Fraassen, and Frederick Suppe, respectively. In addition to reprints of his own papers, Suppe's recent book (1989) provides a good bibliography and a useful participant's overview of these developments (ch. 1, 'Prologue').

7 I have pursued one line of development in Giere 1994b.

8 For an elaboration of the view that crucial experiments are typically after-the-fact reconstructions, see Brannigan 1981.

9 Paul Thagard (1991) is a vigorous exponent of the view that scientific revolutions are to be understood as reflecting the greater explanatory coherence of the victorious theory.

10 I have criticized probabilistic accounts of human judgment in Giere 1988, chapter 6, and discussed Thagard's coherentist approach in Giere 1989a and Giere 1991.

11 For a justification of the strategy of focusing on individuals, see Giere 1989b.

12 I have argued the virtues of a naturalistic rather than a normative account of human judgment in Giere 1988, especially chapters 1 and 6, and further defended this approach in Giere 1989c.

13 Sequential testing, however, introduces the possibility that which model ends up being chosen is a function of the particular order in which alternatives were considered.

14 In this account I have omitted (i) justification for the lack of prior probabilities in the model; (ii) explicit reference to the overall utilities of the decision maker; and (iii) the need for a supplemental decision strategy, such as satisficing, which justifies the obvious decision strategy in terms of satisficing relative to various expected utilities. These aspects of crucial decisions are discussed in Giere 1983, and 1988, chapter 6.

15 Twentieth-century geology, particularly the 1960s 'revolution' in geology, is fast becoming a standard test case for science studies. Among recent books in which it features in whole or in part are those by Le Grand (1988), Stewart (1990), Thagard (1991), and myself (Giere 1988, ch. 8). The articles are already too numerous to list here. One article (Le Grand 1990), however, deserves mention as it explicitly uses images from the 1960s revolution in geology to argue the case for the importance of visual imagery in science.

16 Marvin (1973, p. 43), for example, reproduces an 1858 engraving by Antonio Snyder showing a much too good fit. Speculation regarding the fit of the two hemispheres seems to have followed shortly upon the production of maps of the New World comparable in detail to those of Europe and Africa. This fact fits Latour's (1986) thesis about the importance of 'centers of calculation' for scientific progress. Direct comparisons of the coastlines become possible for a single observer only after many measurements have been brought together in one place and rendered graphically on a single piece of paper for easy viewing.

17 I suspect that Wegener may have been the first to emphasize such geological congruences, but I am not myself sufficiently familiar with the historical sources to vouch for this suspicion.

18 Herbert Simon (1978) distinguishes 'informational' from 'computational' equivalence for representations. Informational equivalence is a generalization of logical equivalence. Computational equivalence means that the same information can be extracted from the representation with the same computational resources. In terms of this distinction, a linear, digital encoding of an image, as for television transmission, may be informationally equivalent to the reconstructed image on a screen. But these two representations are not computationally equivalent. In particular, an ordinary human would find it physically very difficult, if not physically impossible, to extract particular spatial information from the digital representation. But simply by looking at the pictorial representa-

tion, anyone could easily determine, for example, that the cat is on the mat. I would prefer a slightly different terminology, saying, rather, that the two representations are 'logically' equivalent but not 'cognitively' equivalent. But the fundamental idea is the same.

19 This image is taken from the first (1944, p. 506) edition of Holmes's textbook. A very similar diagram appeared in Holmes 1930.

20 It is an interesting question why a dynamic version of Hess's model incorporating geomagnetic reversals does not appear in the literature until after 1966. My suspicion is that publishing conventions in professional journals like *Nature* at that time favoured diagrams and graphs which presented data, as opposed to those that merely pictured speculative models.

10. Are Pictures Really Necessary? The Case of Sewall Wright's 'Adaptive Landscapes'

MICHAEL RUSE

1. INTRODUCTION

Biologists are remarkably visual people. I have before me a flyer from a major publisher, promoting the new edition of an (apparently) highly successful college text in cell biology, co-authored by (among others) the Nobel laureate David Baltimore (Darnell, Lodish, and Baltimore 1990). The 1,105 pages include no less than 1,050 illustrations; the people asked to flack the book harp on the virtues of the pictures ('I appreciate the use of data and actual micrographs. The artwork, and especially the use of color, is outstanding');[1] and instructors adopting the book as a text get a free set of overhead transparencies, with the opportunity to buy more.

Nor is this love of the pictorial confined to the pedagogical. If you look at the papers that biologists produce, and even more at their books, you find them chock-a-block full of photographs and drawings, of graphs and figures, of maps and of stylized tables. Moreover, thanks to advances in technology – photography, computers, printing – the use of pictures of one sort or another is, if anything, increasing rather than otherwise. Bursting with vibrant coloured photographs, some publications seem to owe as much to Walt Disney as they do to Charles Darwin.

Biological illustration has been around for a long time – plenty of time for the philosophers, whose self-appointed task is the understanding of science, to react to it, delving into its nature and significance. So, let us ask about what they have to say – and the answer, I am afraid, is 'remarkably little.' To the best of my knowledge, the classics of logical