A CASE OF NEGATIVE EVIDENCE

Gene Analysis Upsets Turtle Theory

Researchers have concluded from genetic analyses that a widely accepted theory explaining why green turtles migrate 1250 miles to an island in the middle of the Atlantic Ocean to lay their eggs and then swim back to Brazil is invalid.

According to the popular theory, the turtles started coming to the island more than 40 million years ago, when it was close to the shore of South America, and just kept coming as the island moved farther and farther away. But the new research concluded that the turtles had been using the island for only a few tens of thousands of years.

The now-invalid explanation was advanced 15 years ago by Archie Carr and Patrick J. Coleman. It was a bold idea, based on knowledge that the Atlantic Ocean was born 70 million years ago and began spreading. This gradually increased the isolation of Ascension Island, formed by volcanic activity on the ocean's centerline.

As long as the Atlantic was narrow, according to this hypothesis, remote ancestors of today's green turtles had no trouble reaching the island and laying their eggs in its beaches. They then returned to the shallow waters along the Brazil coast to feed on its marine grasses. After 40 million years, the ocean had grown to substantial width, but, according to the hypothesis, the turtles continued to reach the island by some mysterious form of navigation that still enables them to find it.

Now three scientists have examined the extent to which genetic material in the Ascension Island turtles has changed from that of the same species elsewhere. The difference, they believe, is far too small to have evolved over 40 million years. But it is sufficient to show that the Ascension Island turtles are a distinctive group from those nesting at other Atlantic sites, having probably used the island for not more than a few tens of thousands of years.

A comparison was made of a specific locus (the mitochondrial DNA) in the genetic material of turtles from four widely separated regions of Earth. Earlier work had indicated how fast, in a particular population, subtle changes in this material occur over centuries or millions of years. In the 1987 nesting season, eggs, or turtle hatchlings, were taken from twelve nests on French Frigate Shoal in the Hawaiian Islands, ten on Hutchinson Island off Florida, eight on Aves Island off Venezuela, and sixteen on Ascension Island in the mid-Atlantic.

The turtles in Hawaii had presumably been isolated from the other locations since the Isthmus of Panama formed about 3 million years ago. Their DNA, as expected, was most distinctive and provided an index of how fast turtle DNA becomes modified.

Analysis

Step 1. Real World. The real-world subject matter is the population of turtles that has been found to migrate between the coast of Brazil and Ascension Island 1250 miles out into the middle of the Atlantic Ocean. The question scientists have asked about these turtles is how they ever came to make such a long and difficult trip to lay their eggs.

Step 2. Model. According to the previously accepted "standard" model, the members of this species of turtles began visiting the island roughly 40 million years ago when it was quite near the coastline of Brazil. But over the next 10 to 20 million years, the island moved farther out to sea as the floor of the Atlantic Ocean spread out, moving South America farther and farther from the middle of the ocean. The turtles, in this model, kept going to the island, even though the trip got longer and longer.

Step 3. Prediction. The model predicts that any measurements of the age of this population of turtles should yield an age of roughly 40 million years. According to the standard model, that is how long ago the turtles would have had to begin migrating to Ascension Island.

Step 4. Data. Examination of parts of the DNA of turtle eggs taken in 1987 from four different groups of turtles around the world provides a measure of how long a particular population of turtles has existed as a distinct population. By this method, it was determined that the Ascension Island turtles have existed as a distinct population for only a few tens of thousands of years.

Step 5. Negative Evidence? The data and the prediction clearly disagree. The prediction is that the age of the population should be measured at roughly 40 million years. The data put the age at more like 40 thousand years. That makes the prediction roughly 1000 times greater than the data indicates! The data, therefore, provide clear evidence that the model fails to fit the real world. As the headline says, the gene analysis upsets the turtle theory.

Note that the analysis has only five steps. The program makes it clear that whenever the data and prediction disagree, your analysis will terminate at Step 5. Note also that, strictly speaking, the prediction is not that the turtle population is 40 million years old. Rather, the prediction must be about the kind of data that would be expected if the model fit the world. So the prediction is that the data obtained by examining DNA samples from turtle eggs would be of the type that is known to indicate an age of 40 million years.

A CASE OF POSITIVE EVIDENCE

New View of the Mind Gives Unconscious an Expanded Role

For decades, mainstream research psychologists suppressed the notion that crucial mental activity could take place unconsciously. But now, in the wake of exciting new studies, experimental psychologists are taking the unconscious more seriously. Among the most influential of the new studies are investigations into the role of the unconscious in the visual perception of objects and words.

One of the main researchers in this new area is Dr. Anthony Marcel of Cambridge University. He has developed a model of unconscious perception in which the unconscious mind perceives and remembers things of which the conscious mind is unaware. One of the most impressive tests of this model involves what is called "unconscious reading."

In these experiments, Dr. Marcel flashes a word on a screen for a very short time. In addition, the word of interest is "masked" by being surrounded with other nonsense words such as *esnesnon*. When asked directly, the subjects were unable to say what real word appeared on the screen. Dr. Marcel then asked his subjects to guess which of two words looks like the masked word. For example, the masked word might be *blood* and the two choices for look-alikes might be *flood* and *week*. The subjects were correct in their guesses an astonishing 90 percent of the time.

Analysis

Step 1. Real World. The aspect of the real world being investigated is the human mind, particularly human perceptual abilities.

Step 2. Model. The model is not described in much detail. The important thing is that, in Dr. Marcel's model, the human mind has an unconscious component that "perceives and remembers things of which the conscious mind is unaware."

Step 3. Prediction. The model, as reported, does not yield any very precise predictions. It suggests, however, that the subjects might be able unconsciously to perceive and remember the masked words that they could not consciously identify. If so, then it would be expected that their guesses about the word should be right more often than not.

Step 4. Data. The data come from the "unconscious reading" experiments. In these experiments, the subjects "guessed" the correct word "an astonishing 90 percent of the time," even though they could not consciously report what word they had seen flashed on the screen.

Step 5. Negative Evidence? No. The prediction and the data clearly agree. The subjects guessed the correct word far more often than not.

Step 6. Positive Evidence? That the data are described as "astonishing" is a good indication that these sorts of results are thought not to be very likely in the absence of something like unconscious perception. An obvious alternative model is that the subjects were just guessing. It is very unlikely, however, that mere guessing between two alternatives would produce 90 percent correct answers. So the data do support the hypothesis that a model including unconscious perception does fit the human mind.

In this example, the model is only vaguely described. It is difficult to see a clear connection between the model and the prediction. Such cases often result in an inconclusive evaluation. What saves this case, however, is that the data are quite dramatic. It is difficult to imagine another plausible model that could explain the data.

A Case of Inconclusive Data

Was That a Greenhouse Effect? It Depends on Your Theory

The memorably uncomfortable summer of 1988 left many Americans with a suspicion that nature is at last getting even for mankind's wanton pollution of the atmosphere. From California to the Carolinas, the summer's heat wave and drought took a sobering toll. Electric power faltered, vast forests went up in flames, river navigation was throttled, and crops failed.

The "greenhouse effect"—the trapping of solar heat by pollutant gases in the atmosphere—became a household phrase. Some climatologists warned that unless we quickly mend our ways, the world's grain belts will turn to dust bowls, coastal regions will be flooded, forests will die, and countless species will become permanently extinct. On June 23, Dr. James A. Hansen of the National Aeronautics and Space Administration caught the nation's attention when he told a Senate committee that the warming trend almost certainly stems from the greenhouse effect. A crisis, he warned, may not be long in coming.

But forecasting climate has never been as straightforward as scientists could wish. Many are not even sure that the summer's weather was really symptomatic of any trend at all. A. James Wagner, an analyst at the Weather Service, acknowledges that during this decade the world has seen the four warmest years of the past century—1980, 1983, 1986, and 1987. "But I do not feel that the evidence is overpowering that this is anything more than a normal fluctuation," he said.

Climatologists have invented several models in an attempt to understand fluctuations in the weather. One such model, which seems to mimic the real climate quite realistically, was devised by Dr. Edward Lorenz of the Massachusetts Institute of Technology. This model, which does not take carbon dioxide into account but does reckon on the interactions of the atmosphere with the ocean, exhibits large variations.

"The Lorenz model was run backward on a computer for the equivalent of about 400 years," Mr. Wagner said, "and the large fluctuations it sometimes produced, which were not entirely random but were not cyclical either, were quite startling." The swings, he said, were as much as ±3.6°F in global temperature from one year to the next. The model sometimes produced clusters in which several years close together were unusually hot—a pattern imitating the real climate of the 1980s.

Analysis

Step 1. Real World. The real-world object of study is the climate of Earth.

Step 2. Model. The model is the "greenhouse model." According to this model, Earth's atmosphere acts like the windows in a greenhouse, trapping light and heat under the atmosphere. Carbon dioxide increases the efficiency with which the atmosphere traps heat.

Step 3. Prediction. The prediction is that Earth's temperatures should be increasing. However, the model is not developed in sufficient detail to permit precise predictions of how much the temperature should increase by specified dates or whether the increases would be uniform around the world.

Step 4. Data. The data include the drought that covered much of the United States during the summer of 1988. Also, the decade of the 1980s contained "the four warmest years of the past century—1980, 1983, 1986, and 1987."

Step 5. Negative Evidence? No. The data and the prediction seem to agree. The summer of 1988 was unusually warm, and for most of this decade, at least in the United States, the weather was quite warm relative to earlier years.

Step 6. Positive Evidence? The data, however, seem to be relatively likely to have occurred even in the absence of any "greenhouse effect." Some climatologists have even developed an alternative model, the Lorenz model, which predicts relatively large fluctuations in Earth's temperature. These past few years could be one of those fluctuations. The data, therefore, are inconclusive regarding the applicability of the greenhouse model to our current climate. That does not mean, however, that the greenhouse effect is not operative or that it will not show up with more dramatic force in the future.

It is important to note that the ±3.6°F change in temperature from year to year is a prediction and does not count as data even though data are often reported as numbers. This is because no physical interaction with the real-world object of study (the temperature of Earth) occurred in order to obtain this number. Rather, a series of mathematical equations were solved by a computer in order to produce the ±3.6°F.

2.11 Crucial Experiments

One special experimental situation has fascinated scientists since the seventeenth century. This is an experimental setup that allows us to make a clear choice between two rival models. If the experiment is well designed, one model will come out the winner and the other the loser. The fate of both is settled in one stroke. Following the terminology of the seventeenth-century scientist, philosopher, and statesman Francis Bacon, experiments of this type are called **crucial experiments.** So powerful are crucial experiments that scientists, and even historians of science, often reconstruct the past to make it seem that a particular experiment was a crucial experiment—even though, in fact, no one at the time thought of it that way!

The Structure of Crucial Experiments

Figure 2.13 provides a schematic picture of a crucial experiment. Some bit of the real world is put through the apparatus, which produces a reading on a scale. The reading is the data. There are two rival models of the material under investigation. For the moment,

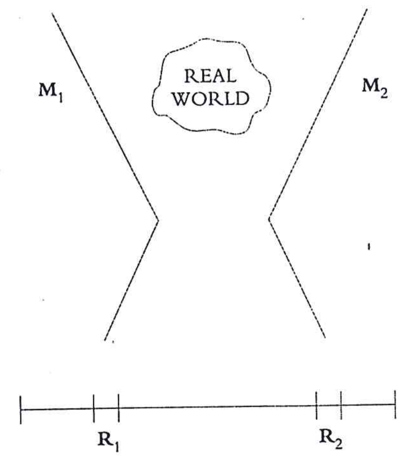

FIGURE 2.13

A schematic representation of a crucial experiment.

we will simply call these M_1 and M_2, respectively. The experiment is cleverly designed so that if M_1 fits, it is very likely that the apparatus would produce a reading in region R_1 of the scale and very unlikely that it would produce a reading in region R_2 of the scale. Similarly, the design ensures that if M_2 fits, it is very likely that the experiment would yield a reading in the region R_2 of the scale and very unlikely it would produce a reading in region R_1. So M_1 predicts a reading in region R_1, and M_2 predicts a reading in region R_2.

The beauty of this design comes across when we consider that the result of the experiment is to be used to decide which model best fits the world. The strategy is that if the experiment produces a reading in region R_1, then we choose M_1 as the best-fitting model. Likewise, if the reading is in region R_2, we choose M_2 as the best-fitting model.

To see why this is an effective strategy, consider the matrix shown in Figure 2.14. Because there are two models to choose between and two different choices that can be

	M_1 Best Fits the Real World	M_2 Best Fits the Real World
Choose M_1 as the Best-Fitting Model	CORRECT	INCORRECT
Choose M_2 as the Best-Fitting Model	INCORRECT	CORRECT

FIGURE 2.14

A matrix representing the choice between two theoretical models in a crucial experiment.

made, the whole experiment has four, not just two, possible final results. Two of these results are *correct* choices: choosing M_1 as the best-fitting model when, in fact, M_1 is the best-fitting model, or choosing M_2 as the best-fitting model when, in fact, M_2 is the best-fitting model. The other two possible results are clearly *incorrect:* choosing M_1 as the best-fitting model when, in fact, M_2 fits best, or choosing M_2 as the best-fitting model when, in fact, M_1 fits best.

The experiment is designed with the presumption that one of the models provides an adequate fit and the other does not. So, initially, there are two possibilities. We can consider each in turn. Suppose, first, that M_1, in fact, provides the better fit. The design ensures that the likely result of the experiment will be a reading in region R_1. In this case, we correctly choose M_1 as the best-fitting model. On the other hand, suppose that M_2, in fact, provides the better fit. The design ensures that the likely result of the experiment will be a reading in region R_2. In this case, we correctly choose M_2 as the best-fitting model. Either way, we are very likely to make a correct choice and very unlikely to make an incorrect choice. The only open question is which of the two possible correct choices we will make. That depends on the data.

In actual practice, experiments that are designed to be crucial experiments often do not work out. The data obtained turn out to be a reading in neither of the regions predicted by the two models. Unless there was some mistake in the execution of the experiment or in the derivation of the predicted readings, we can only conclude that neither of the two models adequately represents the real-world material. The initial presumption that one of the two models fits has to be rejected.

Even though definitive crucial experiments are relatively rare, the possibility of devising such experiments provides a clear, even though idealized, model of scientific reasoning at its best. Understanding this type of scientific reasoning can help us evaluate the less than ideal cases we meet every day.

ANALYSIS OF A CRUCIAL EXPERIMENT

You should attempt to evaluate the following report before reading the sample analysis that follows.

Mutations

It has long been known that changes in the genes of organisms can occur. Such changes are commonly called "mutations." In the 1940s, it was not yet known how mutations occur. Part of the answer to this question was given by a famous experiment performed by Max Delbrück and Salvador Luria. It had recently been learned that some types of viruses (bacteriophages, or phages) can attack and kill some types of bacteria. It is relatively easy to "grow" bacteria in covered dishes containing nourishment in which bacteria generally thrive. These are called "bacteria cultures."

Delbrück and Luria discovered that, in some bacteria cultures, a few of the bacteria survive attacks by phage viruses. Moreover, descendants of the surviving bacteria tend also to survive phage attacks. This shows that the genes of some of the bacteria had undergone mutations that made them resistant to the phage virus and that these resistant bacteria passed their mutant genes on to their offspring.

The question remained as to whether the mutations that made the bacteria resistant were caused by the attacking virus itself or whether they merely happened by chance. The experiment at issue was designed to answer this question regarding the cause of the mutations. Delbrück and Luria considered what would happen if a number (say, twenty) of bacteria cultures, each with a similar small number of bacteria, were allowed to grow for a short time, then injected with the same quantity of phage virus, and then allowed to grow some more. If the phages were producing the mutations, they argued, then all the bacteria cultures should end up with roughly the same number of resistant bacteria.

However, if the mutations were arising by chance, it follows that those bacteria cultures in which the chance mutation happened to occur early in the experiment would end up with many more mutant bacteria than those cultures in which the mutation happened to occur late in the experiment. The earlier mutant bacteria would have a longer time to multiply. Those cultures in which the chance mutation happened to occur at some intermediate time would end up with an intermediate number of mutant bacteria. If it is a matter of pure chance when the mutation occurs, one would, therefore, expect that, by the end of the experiment, there would be a large variation in the numbers of mutant bacteria in the different bacteria cultures.

Delbrück and Luria prepared some bacteria cultures, then introduced the phage virus, and later found that the actual number of resistant bacteria differed widely from one bacteria culture to the next.

Analysis

One way to evaluate this report would be to run through our standard program twice, once for each model. It is quicker and more enlightening, however, to do just one analysis. In this case, many of the steps have two parts.

Step 1. Real World. The real-world object of study is the process by which mutations arise in bacteria.

Step 2. Models. There are two models: (1) according to the *causal* model, the mutations are caused by the action of the phage virus on the genes of the bacteria; and (2) according to the *chance* model, the mutations in the genes of the bacteria occur by chance, independently of any action by the phage viruses.

Step 3. Predictions. There are two predictions, one for each model: (1) the *causal* hypothesis yields the prediction that the number of resistant bacteria will be roughly the same in all the bacteria cultures, and (2) the *chance* hypothesis yields the prediction that there will be a large variation in the numbers of mutant bacteria in the different bacteria cultures.

Step 4. Data. They found that the "actual number of resistant bacteria differed widely from one bacteria culture to the next."

Step 5. Negative Evidence? The data agree with the prediction derived from the chance hypothesis. They disagree with the prediction derived from the causal hypothesis. The data, therefore, provide good evidence that the causal hypothesis is mistaken.

Step 6. Positive Evidence? The only rival to the chance hypothesis mentioned in the report is the causal hypothesis, which yields a different prediction from the chance hypothesis. It is difficult even to imagine another hypothesis. So it does not appear likely that there would have been a large variation in the number of resistant bacteria if the chance hypothesis were false. The data, therefore, provide good evidence that the chance hypothesis is correct.

2.12 MODEL DEVELOPMENT

We need to address the instance in which a well-accepted model is not supported by an experimental test and then is adjusted in some way in order to account for the negative evidence. We will call these instances **model development.** There are times in which the old model will be adjusted slightly and other times in which the old model will be altered dramatically or even replaced with a very new way of thinking about the world. Regardless, the sequence of events is very different from the instance in which two independent models are competing with one another, as in a crucial experiment. Instead, we have a process of developing models in which the two models are dependent upon one another in the sense that one is derived from the other or one follows the other in a sequence of testing and developing models. You will find many articles in the popular press that deal with model development, since an old model's fall and a new model's rise are very newsworthy events in the process of scientific research.

Model development takes two forms: *model development/no experimental test of new model* and *model development/experimental test of new model.*

For the instance of model development/no experimental test of new model, you will find the following sequence of events in the article:

a. An old model is presented.
b. An experimental test of the old model yields negative evidence.
c. A new model is formulated that accounts for the negative evidence obtained for the old model.

The analysis of this type of article utilizes the following steps of the six-step program:

Old Model

1.
2.
3.
4.
5.

A complete analysis of the new model is not possible since no experiment was designed to test the new model.

For the instance of model development/experimental test of new model, the new model is tested experimentally, thus yielding the following sequence of events that is outlined in the article:

a. An old model is presented.
b. An experimental test of the old model yields negative evidence.
c. A new model is formulated that accounts for the negative evidence obtained for the old model.
d. The results of an experimental test of the new model are reported.

The analysis of this type of article utilizes one of the following two sequences of steps:

Agreement between Data and Prediction for the New Model

Old Model	*New Model*
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.
	6.

Disagreement between Data and Prediction for the New Model

Old Model	*New Model*
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.

The following press report illustrates the case of model development.

A CASE OF MODEL DEVELOPMENT

Life on Earth Came Earlier, Crystal Evidence Suggests.

Earth may have been cool enough to retain oceans and develop continents 4.4 billion years ago, according to a new study, possibly pushing back by several million years the time when our planet could have supported life.

The accepted view of an infant Earth is that shortly after it first formed 4.5 billion years ago, the planet became little more than a swirling ball of molten metal and rock. Scientists believed it took a long time, perhaps 700 million years (0.7 billion years), for the Earth to cool to the point that oceans could condense to form a thick, Venus-like atmosphere.

An international group of scientists probed a 4.4 billion-year-old single tiny grain of zircon, a mineral commonly used to determine the age of rocks. The chemistry of the mineral and the rock in which it developed could only have been formed in a low-temperature environment on Earth's surface, the researchers say.

The new evidence led scientists to revise their thinking of how the early Earth formed. Now scientists think that at about 100 million years (0.1 billion years) after Earth gathered itself together from the leftovers of the Sun's formation, temperatures cooled to near or below the boiling point of water.

"This appears to be evidence of the earliest existence of liquid water on our planet," said Margaret Leinen, Assistant Director of Geosciences at the National Science Foundation, which helped to fund the study. "If water occurred this early in the evolution of the Earth, it is possible that primitive life, too, occurred at this time."

Analysis

Step 1. Real World. The aspect of the real world that is the focus of study is the rate of cooling of the early Earth.

Step 2. Model. The old model, referred to as the "accepted view of an infant Earth" in the article, is that "shortly after it first formed 4.5 billion years ago, the planet became little more than a swirling ball of molten metal and rock. Scientists believed it took a long time, perhaps 700 million years (0.7 billion years), for Earth to cool to the point that oceans could condense from a thick, Venus-like atmosphere."

Step 3. Prediction. By examining the chemistry of a grain of zircon that was formed 4.4 billion years ago, one should find evidence of temperatures that were much too high for water to condense.

Step 4. Data. "An international group of scientists probed a 4.4 billion-year-old single tiny grain of zircon, a mineral commonly used to determine the age of rocks. The chemistry of the mineral and the rock in which it developed could only have been formed in a low-temperature environment on the Earth's surface, the researchers say."

Step 5. Negative Evidence? The data do not agree with the prediction. Thus, the data provide evidence that the old model does not fit the real world.

The new model, described in the article, accounts for the negative evidence obtained for the old model by claiming that "about 100 million years (0.1 billion years) after Earth gathered itself together from the leftovers of the Sun's formation, temperatures cooled to near or below the boiling point of water." It is important to note that the apparent agreement between this new model and the data does not count as evidence in favor of the new model. By explicit design, the new model was designed to do just that. The negative evidence for the old model caused scientists to wonder if another explanation might work better. A new experimental test, designed to test the new model, is now needed in order to determine how well the new model fits the real world.

It is tempting to analyze a model development article as a crucial experiment. This results from the fact that model development articles describe an instance in which the data seem to agree with the prediction for one model (the new model) and the data disagree with the prediction for another model (the old model). In model development, the negative evidence for the old model leads to the creation of the new model and is not a genuine experimental test of the new model. In the example given, it is clear that the experiment described provided negative evidence for the old model and that the new hypothesis was developed in order to account for the negative evidence for the old model. If a contest analogy is used to describe a crucial experiment in which one model wins and the other model loses, the analysis of a model development article as a crucial experiment means that the contest is very unfair because the new model was formulated to account for the negative evidence (i.e., the contest is rigged!).

EXERCISES

Analyze these reports following the six-point program for evaluating theoretical hypotheses developed in the text. Number and label your steps. Be as clear and concise as you can, keeping in mind that you must say enough to demonstrate that you do know what you are talking about. A simple "yes" or "no" is never a sufficient answer. Some of these reports are taken directly from recent magazine or newspaper articles and are presented here with only minor editing.

EXERCISE 2.1

Einstein's Impossible Ring: Found

A phenomenon first predicted by Albert Einstein in 1936, and then dismissed by him as something that would be hopeless to look for, has now been found. Astronomers conducting a survey of radio sources at the Very Large Array [VLA] radio telescope near Socorro, New Mexico, have discovered an object in the constellation Leo that has been imaged into a complete ring by the gravitational lensing effect.

"Of course, we'd all heard of Einstein rings," says team member Jacqueline Hewitt of the Haystack Observatory in Massachusetts. "But when I saw it come up on the computer screen, I thought at first it was a problem with the [VLA's image analysis] software."

It was not. Yet her skepticism was understandable. According to Einstein's general theory of relativity, gravitational lensing would happen when light or radio waves from a distant galaxy or quasar pass by a massive foreground object on the way to Earth. The object's gravity would deflect the radiation and thus produce one or more distorted images of the source. A number of such images have actually been found during the past decade. As Einstein himself pointed out, however, the image can only form a complete ring if a source and the lensing object are precisely lined up with Earth—which seems absurdly improbable.

Except that there it was on Hewitt's computer screen—radio source MG1131 + 0456, a tiny oval about 2 arc seconds across with elongated bright spots at either end. In subsequent observations, Hewitt and her colleagues were able to rule out the possibility of its being a supernova remnant or any other such ringlike structure. Moreover, using a regular optical telescope, they obtained optical images of a candidate for the imaging mass: a 22nd magnitude object whose shape and other characteristics are those of a large elliptical galaxy.

EXERCISE 2.2

Why Is the World Full of Large Females?

Animals come in a vast range of sizes, from the tiniest zooplankton to the largest whale. Absolute body size has a crucial influence on a species' life history, affecting such factors as metabolic rate, longevity, and territorial range. And, within a species, relative body size—females compared with males—is important in behavioral ecology terms, too. In most species in the world, females are larger than males, although this rule applies more to groups such as insects, fishes, amphibians, and reptiles than it does to mammals and birds. Nevertheless, the largest animal that has ever lived is a female: the female blue whale.

Why females should attain a larger body size than males has long fascinated biologists. Darwin had an explanation for it, namely: "Increased size must be in some manner of more importance to the females . . . and this perhaps is to allow the production of a vast number of ova." This so-called fecundity-advantage model "has achieved the status of conventional wisdom," says Richard Shine of the University of Sydney, Australia. The model appeals through its simplicity and its consistency with many empirical observations. However, it has not been formally tested, says Shine, a deficiency he has recently repaired. He finds that even though the model may apply in some species, it is by no means universal.

It is no easy task, of course, to solve the question of why one sex may be bigger than the other in a particular species, not least because there are two partners in the game. Specifically, the female might be the bigger of the sexes because of the kind of selective advantage that Darwin proposed, but it is equally true that if males evolve small body size for some different adaptive reason, then the same pattern of body size dimorphism would apply. Several biological factors are likely to be operating in any particular case, and this should always be borne in mind when looking for *the* factor.

Shine elected to test the model in something of a roundabout fashion, thereby hoping to avoid confounding variables that might affect body size in different directions. He measured male-female body size differences in a series of lizard species, some of which produce variable clutch sizes while in others the clutch size is constant. "If the main selective pressure for large female size is an associated increase in fecundity," says Shine, "the species with invariant clutch sizes would have no such advantage and females should tend to be smaller (relative to males)."

It turns out that in anoline iguanids, which produce a single egg, the proportion of species in which the female is larger than the male is about the same as in other iguanids in which clutch size is variable. "The same tends to be true for other lizards with invariant clutch sizes," says Shine. "These data, involving at least seven separate phylogenetic lineages of lizards, appear to falsify the main prediction of the fecundity-advantage model."

EXERCISE 2.3

Prions

In the mid-1970s, biologists generally believed that infectious agents must contain genetic material made of nucleic acid (DNA or RNA) to multiply in their victims. Over the next decade, this belief was challenged by research into the causes of rare degenerative diseases of the brain called "spongiform encephalopathies" because they produce holes in the brain, leaving it looking somewhat like a sponge. The most common such disease, which occurs in sheep and goats, is called "scrapie" because its victims sometimes become so disoriented that they scrape off pieces of their own wool before they die. The most well-known recent outbreak of these diseases was an epidemic of "mad cow disease" in Great Britain in the late 1980s and early 1990s, an outbreak traced to the existence of ground-up sheep heads in commercial feeds. There are four known *human* varieties of the disease, including one, Kuru (laughing death), which existed only among a tribe in Papua, New Guinea, and was apparently transmitted by the (now discontinued) practice of eating the brains of one's deceased relatives.

Suspicions that scrapie was transmitted by proteins alone, without accompanying genetic material, arose first among researchers in Britain. Extracts from the brains of scrapie victims were subjected to ionizing radiation and injected into the brains of normal animals. Ionizing radiation should break down any DNA or RNA present in the extract. Yet these irradiated extracts produced scrapie in the previously healthy experimental animals. Other researchers then subjected similar extracts to procedures known to break up proteins. These treated extracts had a greatly reduced ability to produce scrapie in healthy animals. Some researchers in both Britain and the United States concluded that, somehow, proteins alone were able to produce scrapie in healthy animals. An American researcher, Stanley Prusiner, suggested the name *prions* for these infectious agents to distinguish them from agents such as viruses and bacteria. [For more information, see the articles "Prions," *Scientific American,* October 1984; and "The Prion Diseases," *Scientific American,* January 1995, both by Stanley B. Prusiner.]

EXERCISE 2.4

Identifying the Scrapie Prion

Having shown that proteins can produce scrapie, the next task was to isolate the particular protein that produces scrapie. This search led to a particular protein named PrP, for "prion protein." Surprisingly, however, PrP turned out to be produced naturally in a great many mammals, including mice, with no harmful effects whatsoever. Had they made a mistake?

Was PrP not the infectious agent after all? Prusiner suggested that there might be two forms of PrP, one common and harmless, the other rare and deadly. Having identified PrP, it became possible to isolate the particular genes responsible for its production. Moreover, it was discovered that victims of scrapie and related diseases had a slightly mutated version of the gene that produces normal PrP. With these discoveries, it became possible selectively to breed mice that produced normal PrP, others that produced low levels of the deadly PrP, and still others that produced high levels of the deadly version. It was then possible to construct a persuasive experiment showing that the mutant PrP can indeed transmit scrapie-like infections from one animal to another.

Mice producing high levels of the deadly PrP start out healthy but eventually die of a scrapie-like disease. Those producing low levels of the deadly PrP remain healthy. However, if extracts from the brains of a dead high-level producer are injected into the brain of a low-level producer, it too dies of the disease. If, moreover, extracts from the brain of this second victim are injected into still another low-level producer, it too dies of the disease. Without the injections, neither of the low-level producers would have contracted the disease.

EXERCISE 2.5

HIV versus the Human Immune System

One of the many mysteries surrounding the human immunodeficiency virus (HIV), the cause of AIDS (Acquired Immuno-Deficiency Syndrome), has been how, despite great variability, it almost always manages to defeat the immune system of its host. Some victims succumb within a few years of infection. Many others survive with few symptoms for as many as 10 years before the disease takes over. A few have survived without overt symptoms for as long as 15 years. How can this be?

One hypothesis, described in the August 1995 issue of *Scientific American,* is that the population of viruses actually evolves in the environment of the immune system of its individual victims. The means by which the virus replicates is known to produce relatively large numbers of mutations. While retaining their ability to attack the agents of the immune system, these mutated viruses would become invisible to those agents until they themselves become sufficiently numerous to induce production of new immune system agents designed specifically to attack the new mutants. All the viruses, however, could attack any immune system agent, whereas the immune system agents could only attack those particular mutant viruses that they were designed to recognize. Thus, the forces of the immune system would become more and more specialized and divided. The total size of the virus population would continue to increase. Eventually, the immune system gets overwhelmed, and the virus population takes over. Evolution, however, depends on chance mutations and the environment (in this case the immune system), which varies greatly from individual to individual. So one would expect a wide variation in the length of time it takes for the virus to take over.

The genetic structure of the viruses can now be ascertained in the laboratory. It is thus possible to assess the genetic variability of the virus population in an individual victim as the disease progresses. For the few patients for which this has been done, the results follow the suggested evolutionary pattern. Immediately after infection, there are few variants of the virus. Later, before there are manifest symptoms of the disease, more varieties of mutated viruses emerge. Finally, when the disease becomes manifest, a few variant viruses predominate.

EXERCISE 2.6

A Heresy in Evolutionary Biology

As anyone with even a passing knowledge of evolutionary biology knows, natural selection is a twofold process: the generation of genetic mutations followed by the fixation of variants that are favored by prevailing conditions. And in the world of evolutionary biology, one thing has seemed certain: the generation of genetic mutations is a continuous and random process, uninfluenced by external circumstances. However, if John Cairns, Julie Overbaugh, and Stephen Miller of the Harvard School of Public Health are correct in their interpretation of certain experiments with the bacterium *Escherichia coli,* that certainty may be on shaky ground.

One of the experiments involves taking colonies of *E. coli* that are incapable of metabolizing lactose and exposing them to the sugar. If the lactose-utilizing mutants simply arise spontaneously in the population and are then favored by prevailing conditions, then this would lead to one pattern of new colony growth. A distinctly different pattern is produced if, under the new conditions, the rate of production of lactose-utilizing mutants is enhanced. The observation is something of a mixture of patterns, indicating that directed mutation appears to be occurring. "This experiment suggests that populations of bacteria . . . have some way of producing (or selectively retaining) only the most appropriate mutations," note Cairns and his colleagues. They cite two other types of experiments that can also be interpreted in this way.

Because the randomness of mutation has been so fundamental to evolutionary biology since the 1940s, few researchers have cared to test the notion directly. There are therefore no data beyond those from this handful of experiments that might indicate how general a phenomenon directed mutation might be. Nevertheless, Cairns suspects that it might well turn out to be rather widespread, at least in bacteria. Kent Holsinger, a theoretical population geneticist at the University of Connecticut, says that "if it is general and not just confined to *E. coli* and other bacteria, it could have major implications for evolutionary biology. At the very least, he notes, "there is something going on here that we haven't considered."

EXERCISE 2.7

Discovery Supports Theory That Meteor Caused Dinosaur Extinction

Researchers say they have strong new evidence that the age of dinosaurs ended 65 million years ago when a giant meteorite or comet slammed into Earth with the energy of a billion atomic bombs. Scientists with the U.S. Geological Survey office in Denver said last week that microscopic particles of quartz found in Europe, New Zealand, the Pacific Basin, and elsewhere contain structural cracks associated with the impact of a large body hitting Earth. The mineral debris indicates that a single catastrophic event ended the 150-million-year reign of the great lizards, they said in a new report. The microscopic fracturing found in the quartz is more like that associated with the pressures of a massive impact than what would have resulted from volcanic activity, they say in a study published in the May 8, 1987, issue of the journal *Science.*

Bruce Bohor, Peter Modreski, and Eugene Foord said that the "shocked quartz" is found in the same sediment layers that contain unusually high levels of iridium, a metal common in asteroids, meteors, and comets. The researchers said the latest findings bolster the controversial 10-year-old theory of Nobel Prize-winning physicist Luis Alvarez and his geologist son, Walter, that a single catastrophic event led to a great extinction of life on earth. Geological evidence appears to show a massive extinction of dinosaurs beginning about 65 million years ago, but scientists have been unsure of the reason why.

The Alvarez theory says the impact of the extraterrestrial body released energy equivalent to that of 6 billion Hiroshima atomic bombs and threw up a giant cloud of debris that encircled the globe and diminished sunlight for months, if not years. Climate cooling resulting from the dust blocking sunlight produced the death of dinosaurs and many other types of animal and plant life, according to the theory.

The theory is based on finding up to 600 times normal levels of iridium in clay deposits from the period, and Bohor said in a telephone interview that the same iridium concentrations have been found at every site of the telltale quartz particles. The even distribution of the shock quartz and certain minerals combined with it point to a big comet or meteorite striking a continental area in the Northern Hemisphere, he said. A body 6 miles wide hitting Earth at 45,000 miles per hour, as calculated by the Alvarez theory, could have blasted debris high enough into the atmosphere to account for the worldwide shock quartz distribution, Bohor said.

High pressure associated with volcanic activity can fracture quartz crystals, and proponents of the volcanic theory of dinosaur extinction say this is the source of the shock crystals found in sediments from the period. However, Bohor and his colleagues said quartz fractures caused by impacts are distinct from those resulting from other pressures. "When a meteorite strikes Earth, the mass and speed of impact cause a shock," Bohor said. "This shock wave bounces around in different directions as it hits other objects and comes back into the crystal to produce multiple sets of fracture features unique in impact shocks."

The report said multiple sets of fractures were evident in quartz particles found at several sites: Stevns Klint and Nye Klov in Denmark; Petriccio and Pontedazzo in Italy; Caravaca in Spain; Woodside Creek in New Zealand; and a core taken from the north central Pacific Ocean basin. The samples matched those previously found in Montana and Wyoming, Bohor said, and Soviet scientists recently found the same type of quartz on the east side of the Caspian Sea in the Soviet Republic of Turkmen.

EXERCISE 2.8

The Expanding Universe

One of the most interesting discoveries of the twentieth century is that the universe is expanding (i.e., the galaxies are all moving away from each other). This discovery stimulated the creation of many models of kinds of systems in which such expansion would take place. Two of these models were widely regarded as possibly representing the structure of the real universe. One was an *explosion model* (the "Big Bang" theory) in which all matter is originally concentrated in one place and explodes outward. The other is a *steady-state model,* in which subatomic bits of matter are created out of nothing and eventually move outward, leaving each region of space with the same total amount of matter for all time.

If the universe is an exploding system, it follows that the density of matter (the number of galaxies per cubic light-year) gets less and less the farther away from the original explosion one gets. If the universe is a steady-state system, however, the density of matter should be exactly the same everywhere. This remains true wherever in the universe one happens to be. Whichever direction one looks, the density of the most distant galaxies should be less if the explosion model is correct.

To decide which of these two models best fits the real universe, what we need to do is measure the density of matter in the most distant regions of space. In recent years, radio telescopes have made it possible to make such measurements. These measurements show a clear decrease in the density of the most distant observable galaxies.

EXERCISE 2.9

New Observations Reveal Cosmic Mystery

In the 1920s, it was discovered that the stars exist in large spiral or platter-shaped clusters we now call galaxies. Each galaxy contains millions of stars, with most near the center and fewer out toward the edges. Our Sun and solar system are now thought to be roughly one-third the way in from the edge of the galaxy we call the Milky Way.

Since the discovery of galaxies, astronomers and astrophysicists have naturally wondered how they work and, in particular, what keeps them together. The standard idea has been that a galaxy is held together by the force of gravity. If this is right, then it is possible to calculate the motions of various stars within the galaxy. Unfortunately, it has not until now been possible to measure the motions of stars with sufficient accuracy to determine whether the calculated motions are correct or not.

Recently, however, computers have been used to sharpen the images of stars produced by large telescopes. This new technique has revealed that the stars on the outer edge of the Milky Way are moving much faster than they should according to the standard calculations.

The result has scientists baffled. They have not yet been able to come up with an alternative model that might explain this surprising result.

EXERCISE 2.10

Scientists Put a New Twist on Creation of the Universe

A "Great Wall" of galaxies stretching hundreds of millions of light-years across the known universe has been discovered by two Harvard astronomers and threatens long-held fundamental theories of how the universe came into existence.

The wall, some 500 million light-years long, 200 million light-years wide, and 15 million light-years thick, consists of more than 15,000 galaxies, each with billions of stars. Described as the "largest single coherent structure seen so far in nature," the image of the wall emerged after about a decade of effort to map the structure of the universe in three dimensions. Correspondingly, there are large areas of relative void, containing comparatively few galaxies.

The astronomer's findings were published November 17, 1989, in the journal *Science.* Although the area surveyed is huge, it is only a small piece of the known universe. The astronomers involved said the survey area is to the universe what Rhode Island is to Earth, a small percentage. The survey was performed by mapping the "red shifts," or motions, of a large group of galaxies some 200 to 300 light-years from Earth.

The survey, using the 60-inch telescope at the Whipple Observatory in Arizona, was conducted by dividing an area of the universe into thin slices, then noting the position of galaxies within the slices. As more of the slices were completed and layered together, a three-dimensional portrait of the area emerged and slowly revealed the massive wall of galaxies.

If this map of the universe proves to be a reality, it means the long-held theories of a universe evolving from a smooth, super-hot plasma following a Big Bang some 15 billion years ago might be wrong. The concentration of galaxies around voids filled with either mysterious "dark matter" or nothing points to a very uneven, lumpy universe that apparently could not evolve out of the smooth beginnings described by current theories.

What, if anything, is in the voids, remains unknown and could be a key to understanding the universe. Many scientists believe that 90 to 99 percent of the matter in the universe is dark and has yet to be detected. Whether the huge voids, some as large as 150 million light-years in diameter, contain dark matter is not known.

But given their vast size, said Margaret Geller, one of the two Harvard astronomers to make the discovery, "it may make more physical sense to regard the voids as the fundamental large-scale structures of the universe."

The fundamental theory most threatened by the discovery is known as the "standard model." It holds that the universe originated in a "Big Bang" and has expanded in a fairly uniform manner for the past 15 billion years. It calls for a homogeneous universe that is very much the same in every direction. But the new findings raise questions about how such an uneven system of voids and galaxies could have developed out of a uniform, homogeneous beginning in 15 billion years.

"In my opinion, something is fundamentally wrong in how (scientists believe) the structure of the universe formed," she said. "Maybe the universe isn't homogeneous." Geller said the work is important philosophically because "mapping the universe is about finding its origins. If we can't understand where the universe came from, then we can't understand where we came from."

In their paper detailing the discovery of the "Great Wall," Geller and her colleague, John Huchra, hinted that the wall may be much larger than they have measured. The "most sobering" result of their discovery, they said, is that the size of the largest structures they have found "is limited only by the extent of the survey." If the structures and the voids are larger still, then scientists may have to abandon entirely the fundamental idea of a homogeneous universe.

"There is no theory that even comes close to explaining what we are seeing," said Michael Kurtz, also an astronomer at Harvard.

EXERCISE 2.11

Gravity Waves

Newspapers around the world recently carried reports of a striking new experimental test of Einstein's general theory of relativity. According to Einstein's theory, a large rotating mass should give off "gravity waves" that travel at the speed of light. Several scientists have previously claimed to have detected gravity waves, but none of these claims was very well substantiated. The new experiment was made possible by the discovery of a unique "pulsar"—a very dense star that emits powerful radio waves in distinctive pulses. This particular pulsar appears to be rotating in orbit around another large but invisible mass (perhaps a "black hole"). The discovery of this unique pulsar was made using a radio telescope 1000 feet in diameter—the world's largest.

If the observed system indeed fits an Einsteinian model, it should give off gravity waves. Moreover, according to the model, because gravity waves carry off energy from the system, the rate of rotation of the pulsar around its companion body should be slowing down by about one ten-thousandth of a second a year. Using the large radio telescope and "clocks" accurate to 50-millionths of a second a year, the rate of rotation of the system was measured over a period of four years. The most recent measurements show that the system has indeed been slowing down by just about one-ten-thousandth of a second a year—as predicted. Even scientists who had previously doubted the existence of gravity waves had to admit that finding the system to have slowed down by the predicted amount was quite remarkable.

EXERCISE 2.12

The Genetics of Cancer

Most people are familiar with the model of a chromosome as a string of beads, where the beads represent different genes. Since the early 1970s, it has been possible to identify sequences of genes along a chromosome. This led to the discovery that the chromosomes in cancer cells differ from normal cells in their sequences of genes. It was also known that some cancers run in families and thus must be inherited genetically. Yet even though every cell of a person contains a complete copy of all its chromosomes, not all members of such families get cancer, and those who do exhibit it in only a few sites in their bodies. Moreover, most people who get cancer seem not to be members of a cancer-prone family.

One proposal for explaining these diverse findings, first developed in the early 1970s, has recently achieved broad acceptance. The idea, as explained in the March 1995 issue of *Scientific American,* is that it requires more than one change in the genetic structure of a chromosome to make it produce cancer cells rather than normal cells. People in cancer-prone families start off, so to speak, with one strike against them. They are thus more likely than others to accumulate the required set of mutations but still have some chance of never doing so. This proposal became known as the "two-hit" (or "multiple-hit") theory of cancer.

It was not until 20 years later that it became possible to devise a direct test of this proposal. Several researchers at different locations had fortunately preserved samples of brain tumors from patients who suffered repeated onsets in the same site. If the multiple-hit account is right, the cells of the later, more aggressive tumors should exhibit more mutations than the cells of earlier tumors. The earlier mutations, however, should be preserved in the cells of later tumors. To the delight of proponents of the multiple-hit account, that is just what they found.

EXERCISE 2.13

Deducing the Feeding Habits of *Tyrannosaurus rex*

Ever since the first dinosaur fossil was unearthed by Robert Plot in 1676, dinosaur fever has infected almost everyone. Dinosaurs captivate us not only because of the enormous size that some of them attained, but because of their extraordinary diversity. Encompassing such astonishing creatures as the fearsome *Tyrannosaurus rex,* the majestic *Triceratops* and the cunning raptors, dinosaurs are truly fascinating.

From such fascination many questions arise. By looking at fossils, paleontologists have been able to learn much about the dinosaurs. They can infer the appearance of these creatures in lifelike detail as they existed millions of years ago. They can investigate their diet. And they can find out about the way dinosaurs interacted both among themselves and with their environment.

The truth is that dinosaur fossils are rare, much rarer than fossils of most plants and invertebrate animals. How, then, has it been possible to learn so much about these terrible lizards when we have so few bones to study? The answer lies in trace fossils. As the name implies, a trace fossil is a "trace" of an ancient organism such as a footprint, a tooth or bite mark, or a coprolite.

Footprints and tooth marks, yes, but coprolites? A coprolite is a piece of fossilized dung. Why would anyone study fossilized dung? If you really want to know, you would do well to ask Karen Chin, visiting scientist with the U.S. Geological Survey, Menlo Park, California, who is the world's foremost expert on dinosaur dung.

In 1995, Wendy Sloboda and Tim Tokaryk were hunting fossils near the town of Eastend, Saskatchewan. There they discovered a mass 42 cm long, 12 cm high, and 15 cm wide close to the *Tyrannosaurus* skeleton now known as Scotty. In her lab, Chin identified the mass as a coprolite. Though the work was challenging, "everything came out all right in the end." Now, she had to identify the "species feces" question, as she put it. That is, what dinosaur was responsible for the coprolite? Based on the size of the mass, she concluded that it was from a *Tyrannosaurus rex.*

"We've always guessed that *Tyrannosaurus rex* and their cohorts must have been able to crush the bones of the animals they fed on, but now we have the first hard evidence that they actually did," Chin told reporters following publication of her study in the journal *Nature.* Indeed, with the help of Gregory Erickson of Stanford University, she was able to confirm that fragments of bones were in the sample of dung and identified some of them as juvenile *Triceratops* and a juvenile *Edmontosaurus.*

EXERCISE 2.14

Dove Love: Females Lay Eggs on Coo

Oh, the romance of it all: two doves billing and cooing at each other, joining to build a nest for the forthcoming egg.

Time to amend that romantic notion. Yes, the male coos at the female, and scientists had thought he thus stimulated glandular changes in the female that triggered egg laying. It now turns out that it is the female coo that causes follicles in the ovaries to grow and burst to release an egg, according to psychobiologist Mei-Fang Cheng of Rutgers University.

Dr. Cheng devised painless experiments that rendered the female unable to coo. The male cooed as usual, but nothing happened. But when a female heard recordings of her voice, the result was an egg deposited in the glass nesting bowl.

The male's cooing still plays an important role, Cheng cautions. It "is very powerful in inducing the female to call," but her own call is what triggers endocronologic changes, Cheng says.

EXERCISE 2.15

Project

Find a report of the results of some experiment that is relevant to a theoretical hypothesis. You may find an example in a newspaper or news magazine. The Sunday supplement to your local newspaper is a good bet. The weekly science section of *The New York Times* is particularly good. Or you might try some popular sources that specialize in scientific findings, such as *Scientific American, Psychology Today,* or *Science.* When you have found something that you find interesting and substantial enough to work on, analyze the experiment following the standard program for theoretical hypotheses.

This exercise may be turned into a longer project by looking for other sources of information on the same theory. You might be able to uncover a whole history of experiments related to the theory, each of which can be analyzed. You may discover other experiments bearing on hypotheses using similar theoretical models. Or you may discover cases in which such hypotheses were refuted. Perhaps more elaborate models of the same type were then developed to replace the discarded models. You can then look to see whether later evidence supported these new hypotheses, and so on.

One of the side effects of this project is that you may get an idea of the different levels of science reporting in various popular sources. Some sources tell you everything you need to know to evaluate the reported hypotheses. Others give so little information that you cannot tell whether the evidence supports the hypotheses. You are forced to take their word for it. Most sources fall somewhere between these two extremes.