

9

Absence of Evidence

...they are looking in utter darkness for that which has no existence whatsoever.

–Desiderius Erasmus, 1509 (paraphrase)

Not Required to Exist

So far, we have only talked about the various arguments for the existence of God that are based on the common intuition that the universe, as we know it, could not have come about and have the structure it has except by the action of some transcendent reality. We saw that this intuition is expressed in a number of ways, but they all amount to little more than variations on the ancient argument from design. While this argument, as traditionally presented, is well known to be logically flawed, it has reappeared in recent years expressed in the language of science rather than that of philosophy or theology. And it is in scientific language that I have addressed the issue.

I have attempted to show that current scientific knowledge provides a plausible scenario by which the universe, as we observe it, can be self-contained—*not* requiring external creative actions to either come into existence or to maintain that existence. This scenario cannot be proven and may have few, if any, consequences that can be tested by observations. Still, as long as we limit ourselves to established scientific knowledge and discuss its implications, such a discussion remains within the bounds of science.

Those observations we have made, and those theories we have successfully tested, carry with them a great store of credibility just by virtue of their acceptance into the ranks of scientific knowledge. They can be used to make extrapolations into realms beyond the observable and testable. Of course, we must admit the possibility that these interpolations and extrapolations will turn out wrong. However, this fact does not require that we avoid mentioning them altogether. They are informed speculations, the best we can do, and still better than speculating from a basis of no knowledge whatsoever, or worse, abandoning the quest.

The scenario I have presented is consistent with all scientific observations and is

based on scientific theories that are well tested. It is proposed with the modest purpose of providing a counterexample that serves to refute any assertion that the universe and its laws, as revealed by scientific data, cannot be explained without introducing an external, supernatural creator. Again, I do not claim that this proves that such a creator does not exist, just that one is not required to exist according to our best current knowledge. This makes belief in such a creator optional, but an uneconomical option with no basis in science.

In this chapter, we move to a discussion of the down-to-earth observations that are frequently introduced to profess direct evidence for the existence of a deity who plays an active, everyday role in human life. I will show that we can find no such evidence. And, yet again, because I can predict the line of criticism that this book will generate, I need to make it clear up front that I am not claiming that the absence of evidence eliminates all possibilities for a god to exist in every conceivable form. And I am not reviewing all the theological and philosophical arguments for or against God. I am simply evaluating the *scientific* arguments and *scientific* evidence claimed for a deity according to the same criteria that science applies to any extraordinary claim. I conclude that, so far, they fail to meet the test.

Prediction and Prophecy

Although the Bible should properly be viewed as a religious rather than a scientific document, believers often assert that it contains precise predictions, or *prophecies*, which have come to pass. These are then used to testify to the Bible's accuracy as a source of revelatory information about another, spiritual, plane of reality that transcends material reality.¹ I would not even raise the issue of the Bible in this present book of science were it not for the claims of successful predictions comparable to those of science. These ask to be evaluated on a scientific basis.

The ability of science to predict the future is an important quality that testifies to its power and provides a good basis for believing that it correctly deals with at least some aspect of objective reality. The successful prediction by Edmund Halley (d. 1742) that a comet would reappear in 1758 did much to convince people of the validity of Newtonian physics. In 1846 French astronomer Urbain Jean Joseph Leverrier (d. 1877)

accounted for perturbations in the orbit of the planet Uranus by proposing the the existence and position of a new planet. That same year, the German astronomer Johann Gottfried Galle discovered the planet, now called Neptune, within one degree of the predicted position. Einstein's 1915 theory of general relativity predicted the bending of light by the Sun. Observations made shortly afterward by Sir Arthur Eddington during a solar eclipse in Africa seemed to verify that prediction, however, these were ambiguous. Nevertheless, many increasingly accurate observations since have solidly confirmed the validity of Einstein's prediction.

In 1931 Wolfgang Pauli predicted the existence of an elementary particle with zero electric charge and zero or very low mass that Enrico Fermi named the *neutrino*. This particle was observed in the laboratory in 1956 by Frederick Reines and Clyde Cowan. Neutrinos have now been observed from the Sun and from a 1987 supernova explosion in one of the Milky Way's satellite galaxies, the Large Magellanic Cloud.² In the 1970s physicists used the new standard model to predict the existence of fundamental particles called the W- and Z-bosons, which were theorized to account for the weak nuclear forces. These forces provide for the energy of the Sun and make our very existence on Earth possible. The bosons were observed in experiments with exactly the masses and other properties predicted. And as a final example, I have already mentioned the successful prediction of the inflationary big bang model that the cosmic microwave background should have a temperature fluctuation of one part in 100,000 (thirty millionths of a degree Celsius).

These predictions were not just vague generalities, such as "Two nations will go to war and one will win" (Woody Allen, in a takeoff on Nostradamus), or sure things, like "The Sun will rise tomorrow." The scientific predictions I have mentioned were all *risky*. In each of these examples, the failure of the prediction would have falsified some part of the theory. I can give examples of falsified theories as well, such as an attempted extension of the standard model whose prediction that protons would disintegrate at a certain rate was not verified by experiment, thus eliminating that particular theory.³

Early in the twentieth century, philosophers Rudolf Carnap and Karl Popper proposed that the falsifiability of a theory is what defines it as science. While this proposition is no longer regarded as either necessary or sufficient for a theory to be

scientific, falsifiability remains a major component of scientific method. A theory that can be falsified by some empirical test carries much greater weight than one that only explains observations after the fact. A risky prediction of a theory is one that could falsify it, and when the theory passes such a test, it earns greater credibility.

Examples of successful, risky, falsifiable predictions, and I have only mentioned my personal favorites among thousands, provide us with the confidence to view as reasonable those other assertions science makes that may not be immediately testable. They also provide good reason to believe that science deals with the real world, not one of myth and fantasy. This does not mean that all the assertions scientists make will necessarily prove correct, but rather that they should be taken more seriously than a simple, random shot in the dark. Similar success for the prophecies of the Bible, or any other sacred text, would equally well enhance the credibility of its many assertions that may similarly not be amenable to scientific test.

Most of the biblical predictions that theists point to can only be checked within the Bible itself, which makes them self-fulfilling. In particular, hundreds of "incredible" prophecies referred to by Christians have to do with the ultimate appearance of Jesus as the long-awaited Messiah. This prediction and many others can only be checked by referring back to the Bible for verification. Furthermore, these events are being interpreted today, thousands of years after they are supposed to have taken place. The prediction of the reappearance of Halley's comet in 1758 would have carried far less force had it been made by modern astronomers rather than one who died in 1742.

Obviously, an event cannot be said to be predicted when it occurred before the prediction itself was made. It is easy to predict that if we go to the library and pick up a book on recent American history, published in 1964 or later, we will read that President John F. Kennedy was assassinated in November 1963. Note that the prediction was that we would find the reference, not that JFK would be assassinated at some future date. To have found such a reference in a book published, say, in 1962 would have been remarkable (unless it was a book by Lee Harvey Oswald).

On some occasions, one may successfully predict the uncovering of a hitherto unknown fact about an event that occurred prior to its prediction. For example, an archaeologist might predict that ruins of a city destroyed by war will be found if one

digs at a very specific site. This is still a prediction of a future event, namely the discovery of the ruins.

In the case of the Bible, one cannot point to a single risky prediction that came true based on evidence outside the Bible itself. Here are just a few of the prophecies of the Bible of specific events that failed to materialize:⁴

- Gen. 13:15, 15:18, 17:8, 28:13-14. God promises Abraham and his descendants all of the land of Canaan. Four thousand years later, this promise has still to come to pass.
- Isa. 17:1. Damascus is predicted to cease to be a city. In fact, Damascus is one of the oldest continuously inhabited cities.
- Jer. 49:33 predicts that Hazor will become an everlasting wasteland in which humans will never again dwell. The King James Bible says it will become inhabited by dragons. None of this has happened.
- Zech. 10:11. The Nile is predicted to dry up. This has not yet happened.
- Ezek. 29, 30. The land of Egypt will be laid waste by Nebuchadnezzar, all its people killed and rivers dried up. It will remain uninhabited for forty years. This did not happen.
- Matt. 16:28, 23:36, 24:34; Mark 9:1, 13:30; Luke 9:27. Jesus tells his followers that he will return and establish his kingdom within a generation, before the listeners die. This did not happen.

Apologists can always come up with explanations for the failures of biblical predictions. One, often heard, is that the deity keeps himself hidden so that humans can exercise their free choice in believing or not believing. I cannot refute that possibility, though I have no reason to believe it. The fact remains that one does not have good

reason to look to any document as an accurate source of information when that document is filled with provable falsities.

At best, the errors of the Bible can be classified as ambiguous failures. More convincing would be unambiguous successes. For example, suppose, for the sake of argument, that Jesus had predicted that a human would walk on the moon two thousand years in the future. Now, wouldn't that have been something? If he or other biblical figures had made any successful, risky predictions of this type, by whatever means, then we would have a difficult time not taking their teachings seriously. As it stands, we have no basis for doing so.

I have already mentioned the statements in Genesis that imply predictions of scientific observations not yet made at its writing but long since proven false, in particular, the immobility and immutability of Earth and the stars and the Earth existing before the Sun, Moon, and stars. There are also frequent references to the "ends of the Earth" and the "four corners" resting on pillars.⁵ Now, of course, these mistakes are perfectly understandable if we assume the writers were normal humans, since they reflect the knowledge of the time. But if the Bible is in fact the revealed word of God, then this would be more evident if it provided an accurate model of the Earth and the cosmos.

If we examine the Bible as a scientific document, we would have to conclude that it has been long since falsified. Again, I do not say that it is a scientific document, and most believers do not take it as such. However, the highly visible conservative Christian media, in the United States at least, continually claim that the events described in the Bible have been seconded by science. I am responding to that claim. This is not to deny any value that the scriptures of all religions may have as literary or moral texts, nor their obvious impact on human history. But because of the lack of success in predicting the future, I suggest that they tell us far more about ourselves than about the nature of ultimate reality. And if science does not always provide an inspiring, poetic image of humanity that reinforces our feelings of self-worth, then it at least should be far more trustworthy than biblical religion in offering a glimpse of the reality that transcends petty human concerns. Perhaps science can teach us humility and self-reliance, and the need to live our lives within the universe as it really is, not as we might

wish it to be.

Statistical Significance

Before proceeding, I need to discuss an important technical issue that will come up in this chapter and the next. When experimental results are presented in scientific papers, they are always accompanied by some estimate of their "statistical significance." While sophisticated statistical techniques are sometimes used, most often the authors simply present what statisticians call the *p-value*. For example, suppose an effect is reported with a *p-value* of five percent. This means that in a long sequence of identical experiments we would expect to observe an effect as great or greater produced by statistical fluctuations in five percent of the cases. This not the same as "the likelihood [or probability] that the results were due to chance."

In a number of papers and books to which I will refer below, the *p-value* is incorrectly characterized as the "probability against a reported effect being the result of chance." Sometimes "the odds against chance" is presented. Thus, a *p-value* of five percent is reported as the odds against chance being 20 to 1. These characterizations are wrong. In fact, it is always possible to get any observed effect by chance. You simply have to repeat the experiment enough times.

As we will see, scientific journals in various fields have differing standards for the threshold *p-value* they require for publication. In medical journals, this is typically five percent. This implies that on the order of every twentieth published experiment reports a result that is simply a statistical fluctuation. In physics, by contrast, the threshold *p-value* is typically 0.01 percent, which means that only on the order of one in ten thousand published effects are simply chance artifacts.

I do not mean to make too serious an indictment of medical journals. Their purpose is to disseminate information to the medical community that may save lives. To serve that purpose, a certain amount of useless information that is simply the result of statistical artifacts may be acceptable. Still, allowing five percent useless information into the sample probably does more harm than good, as patients spend their money on worthless treatments and doctors are distracted from more effective courses of treatment.

In fields such as physics, no urgency exists in disseminating new results and a higher priority is placed on avoiding statistical artifacts. The physics standard takes into account the number of trials of various sorts that are made in typical physics experiments. The publishing threshold is sufficiently strict so that few statistical artifacts are published, but not so strict that nothing is published. The typical threshold p -value of 0.01 percent represents a thoughtful compromise that allows some noise in while making sure that clear signals are not missed. Ultimately, independent replication is still needed to confirm the result of a lone experiment.,

The Bible Decoded

Most scholarly discourse on the Bible bears on history and theology, not science, and so is beyond the scope of this book. However, one rather bizarre and widely reported recent incident involving the Bible does merit discussion within a scientific context. It offers an interesting, highly specific example of how one can be misled by relying on intuitive, commonsense estimates of the likelihood of chance events rather than sitting down and calculating them properly. This is a mistake that is easily made when one fails to count all the possible ways a result can happen by chance, a problem we have already seen with the cosmological fine-tuning and design arguments in previous chapters.

In 1994 Doron Witztum, Eliyahu Rips, and Yoav Rosenberg reported in the journal *Statistical Science* that they had discovered coded information in 78,064 characters of the Koren version of the Hebrew text of the *Book of Genesis*.⁶ In this approach they were following an ancient Kabbalistic tradition of seeking messages hidden in the Torah. Rips is a prominent mathematician at the Hebrew University in Jerusalem. Witztum is an unaffiliated physicist, and Rosenberg is a computer programmer. Specifically, the names of a preselected set of famous medieval rabbis (as published in the *Encyclopedia of Great Men in Israel*) were found in close proximity to their birth or death dates, which the authors found surprising.

Witztum and collaborators used a procedure called *equidistant letter sequences* (ELS) in which one starts with a given letter in the text and repeatedly skips a fixed number of letters, ignoring spaces and punctuation. For example, if you take the first

letter of the Book of Genesis, T, and skip every forty-nine letters, you will spell out TORH, the Hebrew spelling of Torah, the Jewish Bible. This was discovered fifty years ago by Rabbi H. M. D. Weissmandel, along with some other unexpected correlations. Applying ELS by hand is obviously tedious, but modern computers make it possible to try a vast number of characters with many different skip numbers and to seek out a large variety of correspondences. One finds that the word TORH appears 56,769 times in the Torah by this method.⁷ Rips and his coauthors insisted that their results on the rabbis could not be the result of chance, claiming a statistical significance of 1.6×10^{-5} .

In introducing the article, one of the editors of *Statistical Science*, Robert E. Kass, wrote: "Our referees were baffled: their prior beliefs made them think the Book of Genesis could not possibly contain meaningful references to modern-day individuals, yet when the authors carried out additional analyses and checks the effect persisted."⁸

Trying a similar analysis with 78,064 characters of the Hebrew version of the Book of Isaiah and a Hebrew translation of Leo Tolstoy's *War and Peace*, Rips and coauthors said they found many names in close proximity to birth or death dates, but the results were statistically insignificant compared to Genesis.

In 1997, after spending time with Rips in Israel, journalist Michael Drosnin authored a popular book, called *The Bible Code*, which became an instant best-seller.⁹ The Rips paper is reprinted in the book, making it perhaps the most reprinted scientific article of all time. Drosnin followed a decoding procedure similar to Rips and collaborators, arranging the 304,805 Hebrew letters of the Bible into a large array, skipping punctuation and spaces. However, his method was not as carefully controlled. Instead of looking strictly for information specified beforehand, Drosnin seems to have searched thousands of letter sequences until he found something that looked interesting.

Drosnin writes that he discovered John F. Kennedy associated with "Dallas," and "Oswald" as "name of assassin who will assassinate." "Oswald" is also encoded with "marksman" and "sniper" as well as the description of the assassination: "he will strike in the head, death." The name "Ruby" appears connected with "he will kill the assassin." Drosnin is particularly fascinated with assassinations. In his most dramatic claim, the name of Israeli Prime Minister Yitzhak Rabin was found to be connected with "assassin

will assassinate" prior to the actual event. Drosnin also leads the reader to think he determined that the assassination would happen in the Hebrew year beginning in late 1995, but this "prediction" was actually done after the fact.

On September 1, 1994, Drosnin flew to Israel with a letter reporting the code's prediction. It was given to Rabin, but on November 4, 1995, Rabin was assassinated in Tel Aviv. The assassin's name and the location of the tragic event were *later* discovered by Drosnin in the code.

In his book, Drosnin also asserted that the Bible predicted the assassinations of Anwar Sadat and Robert F. Kennedy. He uncovered mention of both World Wars, the Holocaust, Hiroshima, and the Moon landing: "Man on Moon," "spaceship," "Apollo 13," along with the precise date of the landing. The collision of the comet Shoemaker-Levy with Jupiter was also foretold. Drosnin found reference to the Watergate affair: "Watergate," "Who is he? President, but he was kicked out." "Shakespeare" appears with "presented on stage," "Macbeth," and "Hamlet." "Einstein" appears with "science" and "he overturned present reality." The Oklahoma bombing is encoded in detail, including "his name is Timothy," "McVeigh," "day 19," "on the 9th hour," "in the morning," "he ambushed, he pounced, terror," and "two years from the death of Koresh." No mention is made of the September 11, 2001 destruction of the World Trade Center, which happened after the book was published.

Well, you might ask, if this isn't evidence for the veracity of the Bible, what could be? How can science possibly explain these amazingly accurate prophecies of events that actually came to pass? Easy.

Retired U.S. Defense Department cryptologist Harold Gans had earlier examined the work of Witztum and his collaborators and found their methods sound; Drosnin reports this in his book. However, in June 1997, after *The Bible Code* was published, Gans issued a public statement that included the following:

The book states that the codes in the Torah can be used to predict future events. This is absolutely unfounded. There is no scientific or mathematical basis for such a statement, and the reasoning used to come to such a conclusion in the book is logically flawed. While it is true that

some historical events have been shown to be encoded in the Book of Genesis in certain configurations, it is absolutely not true that every similar configuration of "encoded" words necessarily represents a potential historical event. In fact, quite the opposite is true: most such configurations will be quite random and are expected to occur in any text of sufficient length. Mr. Drosnin states that his "prediction" of the assassination of Prime Minister Rabin is "proof" that the "Bible Code" can be used to predict the future. A single success, regardless of how spectacular, or even several such "successful" predictions proves absolutely nothing unless the predictions are made and evaluated under carefully controlled conditions. Any respectable scientist knows that "anecdotal" evidence never proves anything.¹⁰

Eliyahu Rips, the coauthor of the original study who had consulted with Drosnin, was also unhappy with the book, though he is featured prominently and favorably in the text. In a public statement, Rips wrote, in part:

I do not support Mr. Drosnin's work on the Codes, nor the conclusions he derives. . . . All attempts to extract messages from Torah codes, or to make predictions based on them, are futile and are of no value. This is not only my own opinion, but the opinion of every scientist who has been involved in original serious Codes research.¹¹

Other observers have cast doubt on the whole enterprise--not just Drosnin's popularization, but serious code research as well. Australian computer scientist Brendan McKay and collaborators he assembled from Hebrew University, Rips's own institution, have provided a scholarly critique of the original work that disputes its significance.¹² This critique was published by the same journal, *Statistical Science*, that published the Witztum, Rips, and Rosenberg paper. Speaking as he did previously for the editors in an introduction to the McKay article, Robert Kass laments that the publication of the Witztum paper was misinterpreted as a "stamp of scientific approval

on the work." He notes that he wrote in his introduction to that paper that it was being offered to readers "as a challenging puzzle." Kass embarrassingly admits that none of the three referees was "willing to spend the time and effort required to reanalyze the data carefully and independently" and assured readers that this has now been done with the McKay article. Kass writes that the new study comprises a "careful dissection and analysis of the Witztum method" and that the "explanations are very convincing and, in broad stroke, familiar." Indeed, they are very familiar to any scientist or mathematician who has done statistical analyses. Kass supports the conclusions of McKay and his collaborators that the protocols used in the original study were insufficiently strict and the statistical analysis flawed. The editor closes by saying that "the puzzle has been solved."¹³

In their article, McKay and his collaborators assert that the Witztum and Rips case is "fatally defective" and the "result merely reflects on the choices made in designing their experiment and collecting data from it." In other words, they saw in the data what they wanted to see. Although Rips and his coauthors insisted that their data sample was strictly preselected, they were not strict enough and allowed sufficient wiggle room to get artifactual positive results.¹⁴

As you might expect, Rips and Witztum dispute all this.¹⁵ However, the most damaging case against the whole notion of finding messages hidden in the Bible by playing with letters on the computer is that the same can be done with any sufficiently long text, with the same "astounding" results. McKay and his collaborators have provided numerous, amusing examples of how you can find whatever you want, in English as well as Hebrew.¹⁶ Drosnin had stated, "When my critics find a message about the assassination of a prime minister encrypted in *Moby Dick*, I'll believe them." McKay answered the challenge, using Drosnin's methods on *Moby Dick* to "predict" the assassinations of Indira Ghandi, Martin Luther King, John F. Kennedy, Abraham Lincoln, and Yitzhak Rabin, as well as the death of Princess Diana. Others have joined in the fun. New Mexico physicist Dave Thomas performed an ELS analysis on the King James Version of Genesis and found "Roswell" and UFO" hidden within. He also found "Hitler/Nazi" encoded in Genesis, *War and Peace*, and the 1987 U.S. Supreme Court ruling on *Edwards v. Aguillard*, the creationist decision, referred to in chapter 2, which he

conveniently had on his computer .¹⁷

Brendan McKay also reports, on his Web site, that he has found "Brendan McKay," "Chanukah Candles," and "Jesus the Messiah" in *War and Peace*. "Bill Gates" is ominously encoded in the *Book of Revelation*.

Droshin claims that the probabilities of getting many of his results by chance is extremely low. However, it appears he calculated that probability incorrectly, not counting all the opportunities for chance to produce his observed connections in the billions of trials he makes. This mistake is very common. It is called the "filedrawer effect" in statistics and happens to be one that I have often seen made by enthusiastic physicists who are overly anxious to make a major discovery and become famous. Fortunately, the self-correcting procedures in place in physics and other sciences usually catch these human errors before they are announced.

The Bible Code has been widely reviewed, with mixed results. Bible believers have been understandably delighted, and one can find the book promoted by Jewish and Christian organizations. However, I am not aware of any favorable mention of Droshin's book by mathematics or statistics experts, who are only too conscious of the dangers of being fooled by chance events of low probability when a large number of trials is made. Let me just quote from two reviews by mathematicians that appeared back-to-back in the same issue of *Notices of the American Mathematical Society*.

Allyn Jackson writes that "this book is a series of wild, unfounded claims based on stretching statistical evidence to the breaking point." Droshin has said he is simply a skeptical journalist in search of the facts, but Jackson notes that he "lacks the mathematical and statistical background that would bring some depth to his skepticism. He is deluded by his own ignorance."¹⁸

In a following commentary, mathematics professor and Orthodox rabbi Shlomo Sternberg minces no words, writing that "*The Bible Code* by Michael Droshin exploits a hoax perpetrated by two Israelis, E. Rips and D. Witztum." He points out that the messages found by both the Israeli group and Droshin would collapse even if just a few letters were added to or dropped from the text they used. As Sternberg explains, "Any serious student of the Talmud knows that there are many citations of the Hebrew Bible which indicate a differing text from the one we have. . . . One of the oldest complete

texts of the Bible, the Leningrad codex (from 1009) differs from the Koren version used by Rips and Witztum in forty-one places in Deuteronomy alone. In fact, the spelling in the Hebrew Bible did not become uniformized until the sixteenth century with the advent of a printed version that could provide an identical standard text available at diverse geographical locations."¹⁹

Drosnin shows his ignorance of biblical scholarship, claiming, in direct contradiction to Rabbi Sternberg, that "every Hebrew Bible that now exists is the same letter for letter" .²⁰ Drosnin also exhibits an ignorance of statistics. He tells us, "The odds against Rabin's full name appearing with the prediction of his assassination were at least 3000 to 1." Perhaps. But then he says, "Mathematicians say 100 to 1 is beyond chance. The most rigorous test ever used is 1000 to 1" .²¹ Note that Drosnin has misinterpreted the meaning of the p-value in the way I have explained above.

In fact, no mathematician would ever say that "100 to 1 is beyond chance." Recall, from chapter 4, that mathematician William Dembski, in his desire to prove the existence of intelligent design, uses odds of 10^{150} to 1 to define an information sequence as too complex to have happened by chance (although, recall that even this criterion is inadequate after the fact). Furthermore, as we have seen, the threshold used in physics for the publication of a new result is conventionally set at a p-value of 0.01 percent, far more rigorous than the 1000 to 1, or p-value 0.1 percent, that Drosnin tells us is the "most rigorous ever used."

I have gone into the story of *The Bible Code* in considerable detail (and, I have still left a lot out) because it provides a good lesson in how not to do science--especially when you are in the business of looking for God by scientific means. As we saw in the preceding section, a claimed prediction is not a meaningful prediction unless it is explicitly published before the fact. Drosnin's "prediction" of Rabin's assassination was probably truthfully made before the fact, although, to my knowledge, it was not published until after the fact. However, we also saw that the only predictions that are worth anything are *risky* predictions. Given the violence of the Middle East, Rabin's violent death was not a very risky prediction. Furthermore, Drosnin admitted that some details were later filled in after the assassination took place.

You only need to read *The Bible Code* to see just how disingenuous Drosnin is on

this, his major claim. Throughout the book he talks about how he predicted the correct year of Rabin's assassination, implying this was done prior to his going to Israel to warn the prime minister. Yet Drosnin did not find the year, 5756 on the Hebrew calendar, in association with Rabin's name. Rather, he found it in the vicinity of the words "Holocaust of Israel." This a very tenuous connection, indeed. Any one of a thousand phrases might have given the same association: "Fire in Israel," "Loud noise in Israel," "Sad event in Israel," and so on. Make up your own list. Furthermore, when Drosnin described how this prediction came about, he tells us that he made this discovery "While Israel mourned Rabin," obviously *after* the event .²²

And this is not all that is fishy about Drosnin's claimed prediction. The expression "Assassin that will assassinate" is his personal translation, and he admits he knew no Hebrew until he started this project. As we have seen, he is hardly a Hebrew scholar. The Jerusalem Bible translates the identical Hebrew text as it appears in Deut. 4:42 as "a man who has killed his fellow,"²³ This verse, as translated in the Revised Standard Version, talks about a manslayer killing his neighbor unintentionally. Clearly Drosnin has stretched the facts mightily to say that he predicted the details of the assassination of Yitzhak Rabin in the year it happened, before it happened.

Drosnin makes a number of predictions that, as of this writing, are still in the future, such as an earthquake in Los Angeles in 2010. This would not be a highly unlikely or extraordinary event. A meteor hitting Los Angeles would be a more risky prediction, and I am sure he can find that in the Bible if he looks there or in any other sufficiently long text. As Allyn Jackson notes, Drosnin's prediction of an "atomic holocaust" in Israel in 1995-1996 (the same Hebrew year as Rabin's assassination) did not pan out.²⁴ However, he later found the word "delayed" encoded nearby. Like most fortune-tellers, Drosnin hedges his bets so that he can win no matter what happens.

As for the low probability of chance claimed for their results by the more credible Rips and his collaborators, McKay and his coauthors demonstrated that this probability was grossly underestimated because account was not taken of the wiggle room they allowed in the rules used to conduct the experiment. Recall that statistical significance of 1.6×10^{-5} quoted in the Rips paper was based on a choice of thirty-two "famous rabbis" whose names and other biographical information were found

embedded in the Bible. McKay shows that by removing four not-so-famous ones, the results become statistically insignificant. . We should be highly suspicious of any effect that is highly sensitive to the particular sample used. A good experiment must always contain carefully constructed safeguards to prevent the experimenters from, consciously or unconsciously, selecting a sample that gives their desired results and ignoring those samples that do not--the filedrawer effect mentioned above. If several different samples are used in the study, then the calculation of chance odds to be presented in the final report must include those that fail to give the effects being claimed as well as those that do.

This is the same problem that arises in the interpretation of an anecdotal claim for some paranormal event, such as a prophetic dream. To calculate the odds that the event was a chance occurrence, we need to know all the times the same dream occurred without coming true. Few people even remember having such dreams.

I have only presented the first round of arguments on both sides of the Bible code debate. The interested reader can follow up on the responses that the various antagonists have made to the other side's rhetoric. However, I doubt that the Bible side will win this one--unless they discover some new, risky prediction, publish it in a respectable scientific journal before the fact, and then have it come true. Based on the situation, as it now exists, I think we can safely draw the same conclusion we did in the preceding section, that no case can be made for the Bible as a source of revelatory information.

In November, 2002, a sequel Bible Code II: The Countdown by Michael Drosnin was released by Viking Press predicting a range of catastrophes leading to Armageddon. Dave Thomas has taken the 6,966 characters from book excerpts on the Amazon.com web site and extracted this secret message: "The Bible Code is silly, dumb, fake, false, evil, nasty, dismal fraud and snake-oil hoax" (to be published).

The whole Bible code saga provides yet another example of, and further insight into, the thinking processes that we saw repeatedly in our previous review of the design arguments for the existence of God. Hugh Ross, Michael Behe, William Dembski, and Bible decoders perceive God or "intelligent design" in their data because that's what they want to see. They are like the people who imagine a man in the moon or the face

of Jesus on a burned tortilla.

Note that, as was the case with the design arguments discussed earlier, we again find a scientific dispute over how to make proper probability calculations. Ross insists that the fine-tuning of the constants of physics is too improbable to have happened by chance or other natural process. But he does not demonstrate this quantitatively. He simply assumes his readers will find his assertions plausible, since most want to believe them anyway. Behe and Dembski make a like argument, that the order we see in biology is too improbable to be natural. But, again, they can only call upon their readers to take their word for it. The statistical analysis of the Bible code shows that what appears on the surface to be highly improbable can in fact be very likely when you select for the desired outcome. Ironically, that's exactly how natural selection works, with the desired outcome being gene survival. Furthermore, even if Ross, Behe, Dembski, and the Bible decoders are correct in their guesses that the probabilities for certain happenings being natural or purposeless are low, they must also demonstrate why the alternative, nonnatural or purposeful probabilities are not even lower.

Testing for God

Let us now move to a quite different arena in which direct, empirical evidence for the reality of the Judeo-Christian-Islamic God has been recently claimed.²⁵ Indeed, as I have emphasized throughout, it is the existence of this God—one who plays a highly active role in everyday human affairs—and not all conceivable gods, that can be tested scientifically. We simply have to look for the results of all the activity of God that may be reasonably anticipated and attempt to analyze these results scientifically. This should be, in principle, no more difficult than looking for effects of environmental pollution on the growth of corn or sending a beam of neutrinos into a block of matter and studying what comes out of that block.

Whatever other properties the God of Jews, Christians, and Muslims may have, he is a God who answers prayers. As Jesus promised in John 16:23-24: "And in that day ye shall ask me nothing. Verily, verily, I say unto you, Whatsoever ye shall ask the Father in my name, he will give it you. Hitherto have ye asked nothing in my name: ask, and ye shall receive, that your joy may be full." With all the billions of prayers

being said every day throughout the world, their effect should be readily measurable.

The task of any empirical research program is to make observations under carefully controlled situations and to try to measure any causal connections between the data and existing or hypothesized theories. Usually, several different explanations for the observations are possible; otherwise, why do the research? Thus, the experiment must be designed to distinguish between the various possibilities. Even when the data seem to support a certain pet theory favored by the experimenters, they cannot immediately conclude that their theory is thereby validated. Other explanations must be ruled out with high probability. And although great care may be taken to eliminate alternative causes, the possibility remains that the result was simply due to chance—a random statistical fluctuation.

Fortunately, the methods of statistical error analysis are now highly developed, and most scientists are trained in their application. Unfortunately, many authors who claim to make scientific statements are not very knowledgeable in these techniques. We saw this in the above discussion of the Bible code. Recall Drosnin's claim that his prediction of the Rabin assassination had an estimated 3000-to-1 odds against being an artifact of chance. This sounds very impressive, and Drosnin asserted that anything above 100 to 1 is "beyond chance," while claiming that 1000 to 1 is the "most rigorous test ever used."

But, as I pointed out, beside Drosnin's misinterpretation of the p -value, these two statements are wrong. Odds of 100-to-one are hardly beyond chance and tests more rigorous than 1000-to-1 are commonplace in many fields.

Consider the following simple example, which I have used before. Suppose you have five coins. It would seem very unlikely that they will all fall heads (or tails) up in any given toss. However, they will do so, on average, every 32 tosses. Let us make it more interesting and toss twelve coins. These will fall heads up on average once every 4,096 times. These odds are even greater than Drosnin's 3000 to 1. Now, suppose instead of looking just for some preselected pattern, like all heads up, you simply look for anything interesting. That could include all tails up, alternating heads and tails, the first six heads and the second six tails, and so on. Obviously the chance odds can then decrease to the point where "something interesting" would be commonplace.

This was the situation we saw in the discussion of the intelligent design arguments in chapter 4. In claiming evidence for design in nature, William Dembski proposes that we look for complex specified information. He defines complex information as a minimum of 500 bits. The chance odds for any given string of 500 bits is one in 10^{150} , if the string is specified in advance. This does not mean that 500 bits of information is impossible to generate. You can easily produce 500 bits of previously *unspecified* information. Just write down any sequence of 150 digits from 0 to 9; it will take you only a few minutes. Or, you can write a string of 500 binary digits: 0s and 1s. The probability for the sequence generated, after the fact, is 100 percent!

So it is not just the number of bits, the complexity of the information, but the specification of a particular bit sequence that makes the information highly unlikely to be the results of chance. After you write down your 150 digits, you would not reproduce that sequence again, with a random number generator, in the age of the universe. However, as we saw, Dembski does not limit his definition of specified information to preselected sequences, allowing the postselection of "interesting" patterns. While he claims to provide a prescription for distinguishing suitable postselected patterns (which he calls "specifications") from unsuitable ones (which he calls "fabrications"), it is not clear that any such distinction can be made.

Drosnin's 3000 to 1 seems very unlikely by chance until you remember that he was not just looking for "Yitzhak Rabin" but for "John F. Kennedy," "Robert F. Kennedy," "Anwar Sadat," "Martin Luther King," and who knows how many other familiar names. He was not doing a single experiment, like tossing twelve coins once. He was, in effect, tossing twelve coins thousands of times in his computer and selecting whatever trials came out in what he regarded, after the fact, as an interesting sequence. And this is exactly how evolution produces the complex sequences of information that constitute living organisms: interesting patterns are chosen from random ones by natural selection.

With this lengthy introduction, we are now in a position to evaluate, in a scientific fashion, whether events that are claimed to be affected by supernatural intervention are sufficiently significant statistically so as to make highly unlikely their occurrence being merely the result of chance.

If you were to rely on the personal testimony of millions, you would conclude that the supernatural does indeed exist. Most believers can cite instances in which they or someone they know have prayed for a particular outcome and had it happen, or had other personal experiences that convince them of the existence of God. They will readily admit that many of their prayers have gone unanswered, but this does not, in most cases, weaken their faith. Surely it would be unreasonable for God to answer every prayer, or even most prayers. He answers the ones he wants to answer. These may not be many in number, but not zero if he is the God of Judaism, Christianity, and Islam. Just a few clear cases of prayers being answered would be sufficient to prove the existence of that God, if no other explanation can be found.

However, we cannot make a definitive statement based on anecdotes or other uncontrolled observations. These have too many variables that can affect the outcome. We have no idea how to take these variables into account in our calculation of chance odds. For example, how unlikely is it to have a specific dream come true? We can calculate that only by knowing all the times that specific dream did not come true, and records of unfulfilled dreams are not kept. Normally, they are not even remembered. It is not possible to say that some happening was highly unlikely to be the result of chance without counting all the outcomes. For example, suppose that the odds against a particular individual winning a lottery are 10 million-to-one. If a million people enter the lottery, then the odds of someone winning are one in ten. With billions of people on the planet, unlikely events will happen every day by chance. That's why we must limit ourselves to controlled experiments or other types of observations in which all the variables are known and we can properly estimate the odds.

Throughout history, and widespread even today, certain select individuals have been reported as having the supernatural power to work miracles. Suffice it to say, none have done so under controlled conditions acceptable to the scientific consensus. Of special note are faith healers who claim to cure in the name of God. Despite thousands of anecdotes, no faith healer has proved his or her ability in scientific trials. Indeed, many have been uncovered as charlatans who prey upon the sick, taking their money while giving them false hopes and sometimes discouraging them from getting proper treatment. Magician James Randi has skillfully exposed many members of this

unsavory group, although they still thrive on religious cable channels and at revival meetings often attended by thousands.²⁶ I cannot understand why religious leaders do not denounce the practice of faith healing and why laws are not passed against it. Surely these shameless individuals do the image of religion no good.

Faith healing, if it were successful, would provide a good case for God. Its lack of success is a good case against. Not only are people robbed of their money and their dignity by unscrupulous faith healers, those who practice faith healing in an honest and pious belief that they are helping others can end up doing more harm than good. For example, pediatricians Seth Asser and Rita Swan have looked at the cases of 172 children who died between 1975 and 1995 whose parents withheld medical care because of reliance on religious faith-healing rituals. They found that 140 fatalities were from conditions for which survival rates with medical care would have exceeded 90 percent. Eighteen more had expected survival rates of greater than 50 percent. All but 3 of the remainder would likely have had some benefit from clinical help. The authors concluded: "When faith healing is used to the exclusion of medical treatment, the number of preventable child fatalities and the associated suffering are substantial and warrant public concern." The authors call for legislative action to protect these innocent victims of fanatical belief.²⁷

Hundreds of published scientific reports can be found that claim measurable effects of religious or spiritual actions. Virtually all are found in the health domain, in which the possible therapeutic effects of prayer and other religious activities are investigated.²⁸

In general, medical research falls into two categories: (1) *epidemiological*, in which health records of a large population are examined for associations between various characteristics and outcomes; and (2) *experimental*, in which a specific intervention is applied to a sample of patients and compared with a control sample of patients picked from the same set but for which the intervention is not applied. In (2), the patients assigned to each sample are selected randomly, and if the experiment is done properly, neither the patients nor the attending medical personnel know who has been assigned where. That is, the experiment is "blind." Let us take a look at some of the reported results in each of these categories.

The Epidemiology of Religion

Epidemiological studies, in general, are notoriously difficult to interpret reliably because of so-called confounding factors. A given study may indicate a correlation between an illness and some factor, but this does not necessarily prove that the factor is the cause (or cure) of the illness. In an amusing example given by psychiatrist Richard Sloan, a study might find that lung cancer is more prevalent among people who carry around matches in their pockets. It would not follow that matches cause lung cancer.²⁹

Perhaps the first scientist to do an epidemiological study on the possible health benefits of religion was Francis Galton, the prominent nineteenth-century anthropologist (and cousin of Darwin). Galton looked at published data on the mean lifetimes of men in upper classes in England, including members of royal houses, clergy, lawyers, medical doctors, military men, and others. He noted that royalty had the shortest mean lifetime, 64.04 years, despite being the most prayed for. As Galton interprets this result:

The sovereigns are literally the shortest lived of all who have the advantage of affluence. The prayer has therefore no efficacy, unless the very questionable hypothesis be raised, that the conditions of royal life may naturally be yet more fatal, and that their influence is partly, though incompletely, neutralized by the effects of public prayers.³⁰

Clergy did better, living on average 69.49 years, but that was reduced to 66.51 years when the more prominent members of the class were singled out. Galton remarks: "Hence the prayers of the clergy for protection against the perils and dangers of the night, for protection during the day, and for recovery from sickness, appear to be futile in result." Galton's sample is too small to draw any firm conclusion except to say that he found no evidence for a strong effect of prayer on the lives of people who would be the most expected to benefit.³¹

Recently we have seen an upsurge in epidemiological studies that attempt to discern a connection between religion and health. In a 1998 review published in *Archives*

of *Family Medicine*, Dale Matthews and five collaborators examined a large number of papers and concluded that "a large proportion of published empirical data suggest that religious commitment may play a beneficial role in preventing mental and physical illness, improving how people cope with mental and physical illness, and facilitating recovery from illness."³²

One of the coauthors of this study, which was partially funded by the Templeton Foundation, was Harold G. Koenig, director of the Center for the Study of Religion/Spirituality and Health located at Duke University. Koenig is also the author of a popular book, *The Healing Power of Faith*, and has appeared on national television promoting the power of religion to heal. In his book he claims to have evidence that people with strong faith are healthier, experience less depression, and live longer.³³

A review of the literature from a more critical perspective has been provided by R. Sloan, E. Bagiella, and T. Powell in the British medical journal *Lancet*.³⁴ They concluded that the published work lacks consistency and is not based on sufficiently large samples of data. Linda Gundersen has also reviewed the subject in *Annals of Internal Medicine* and finds that the conclusions of many of the studies are doubtful.³⁵

By far, the greatest number of these investigations is of the epidemiological variety, in which various religious behaviors are examined for their possible association with health. Churchgoing seems to produce the greatest correlation, but none of these reports have the statistical significance that merit their being regarded as definitive. Furthermore, the studies have been unable to adequately establish a clean causal connection between religion and health. As Sloan and his coauthors explain, when confounding variables are considered, these can explain most if not all of the effects observed. For example, a 1972 study that is often cited as evidence for a positive association between church attendance and health³⁶ was later found, by one of its own authors, to be mistaken due to failure to control for people with reduced mobility. People in poorer health were simply less likely to go to church.³⁷ Incidentally, this is a nice example of science successfully policing itself.

A similar question regarding the effects of confounding variables can be raised with the frequently heard claims of a connection between positive thinking and health. Perhaps such a connection exists, but is it positive thinking that causes good health or

good health that causes positive thinking?

A much ballyhooed recent report has also been misinterpreted by those who chose not to read carefully what the investigators actually wrote. The autobiographies of 180 catholic nuns were studied, and it was found that those who reported positive emotions in childhood lived longer than those who did not, with p -value of 0.1 percent.³⁸ This did not mean that nuns live longer than the rest of us because of supernatural intervention, just that positive-thinking nuns may live longer than negative-thinking nuns.

The best way to extract a single factor for a difference in observations among two samples is for the samples to be equal in every way except for that one factor. Thus, for experimental purposes, groups of religious people need to be compared to groups of nonreligious people, with the two groups suitably identical in every other conceivable respect. This should be possible. Not all atheists indulge in behaviors that are damaging to their health.

It might even turn out that the atheists do better! Researchers must be careful to guard against any bias against that possibility. After all, far fewer prison inmates are atheists in proportion to their representation in the general population, and far more are Christians. But, I must admit, this has a confounding factor, too: parole boards tend to look favorably on convicts who profess to be "born again."

In fact, studies exist that indicate a negative influence of religion on health, although these are not always publicized with the same fanfare as those that give positive effects. As Kevin Courney has discovered, Koenig and his collaborators selected favorable studies in their review, mentioned above, and ignored even some of their own research that gave conclusions opposite to the ones they preferred to see.³⁹ For example, a 1994 Koenig paper notes that "several studies have reported an association between psychiatric disorder and religious affiliation, with the rates of disorder highest among nonmainline Protestant religious groups."⁴⁰ Another Koenig study ignored in his review showed that the likelihood of major depression among highly religious Pentecostals was three times greater than among persons who said they had no religious preference.⁴¹ Other investigators have also shown that the religious are frequently among the lowest in mental health.⁴² Koenig also fails to

mention other studies on the role of religion in child abuse and medical neglect.⁴³

Another recent report indicates that religion may not be so good for you after all. As reported in the *Archives of Internal Medicine*, Bowling Green State University psychologist Kenneth Pargament and his colleagues, which included Koenig, looked at 595 individuals aged fifty-five or older who had been hospitalized between 1996 and 1997. Over 95 percent were conservative or mainline Christians. Jews, Buddhists, or nonbelievers were not studied. Patients who reported feeling alienated from God or believing the devil brought about their illness were associated with up to a 28 percent increase in risk of dying during the two-year follow-up period. None of the other factors considered, such as gender or race, showed any significant difference.⁴⁴ Of course, this could simply be the effect of negative thinking. I mention this only to show that the effect of religion on health is ambiguous at best and that the one-sided reports one usually reads in the media provide no scientific support for a supernatural role in health.

Even if some positive correlation between religion and health can be found in some areas (mental health does not seem to be one of them), religious people, as a group, are less likely than the general population to engage in risky behaviors such as smoking, excessive alcohol consumption, and promiscuous sex. One can conceive of some value of religious behavior without invoking divine interventions as necessary causes.

For years, Herbert Benson has been studying the health benefits of prayer, meditation, and relaxation.⁴⁵ The statistical significance of some of his claimed positive benefits have been questioned.⁴⁶ But even if the reported effects are real, are they spiritual? Do they provide evidence for a spirit realm? Hardly. The ability of the mind to heal or hurt the body is well known. It seems perfectly reasonable and natural that relaxation by any means can have a benefit, for example, by lowering blood pressure and reducing stress. Benson provides no evidence that a Christian prayer works better than a Hindu one or a simple, perfectly secular relaxation exercise.

Intercessory Prayer

Praying for yourself might help you. Or it might harm you. Either outcome is possible, by purely material brain-body interactions with nothing spiritual required. And praying for another person with his or her knowledge might also help, by purely material means, in reducing that person's stress. Or it might hurt by applying more stress. Believe it or not, prayer may be harmful! Scott Walker and his collaborators have reported on the results of remote prayer on the patients of an alcohol treatment facility. Their conclusion: "Compared with a normative group of patients treated at the same facility, participants in the prayer study experienced a delay in drinking reduction. Those who reported. . . that a family member or friend was already praying for them were found to be drinking significantly more at six months than were those who reported being unaware of anyone praying for them."⁴⁷ However, based on the data, we have to conclude that prayer does little or nothing most of the time.

We can safely put aside all the claims for evidence of a God who answers prayers in the set of epidemiological studies published to-date. Nevertheless, no material explanation is likely to be able to account for intercessory prayer done blindly at a distance—should it ever be successfully demonstrated. In another widely reported investigation (the negative reports, like those mentioned above, rarely make the media), cardiologist Randolph Byrd published an article in a regional medical journal claiming evidence that coronary patients benefitted from blind, distant intercessory prayer.⁴⁸ His patient group was comprised of 192 who were prayed for by born-again Christians from a local church. The control group contained 201 patients. It was claimed that this was a double-blind experiment, with subjects randomly assigned to treatment group or control group, and with treatment staff uninformed of the assignments.

As treatments progressed, the patients in both groups were graded in twenty-six medical categories as "good," "intermediate" or "bad." In six of these, the prayed-for group did better at the level of 5 to 7 percent. In one of the six categories, for example, 7 percent of the patients needed fewer antibiotics. But, no significant difference was found in total days in the hospital or death rates, despite intercessory prayer specifically requesting a "rapid recovery."

How significant is the reported improvement? Byrd himself admitted in the

paper: "Even though for [the six seemingly significant] variables the P values were less than 0.05, they could not be considered statistically significant because of the large number of variables examined. I used two methods to overcome this statistical limitation. . .[the] severity score, and multivariant [sic] analysis."⁴⁹

In other words, the results were insignificant. However, like so many others working in this field, Byrd said something quite different to the press. The *Kansas City Star* reported him as stating unequivocally that "the patients who were prayed for just did better."⁵⁰ Has anyone ever told you not to believe everything you read in the papers (or hear on radio, see on television, or read on the Internet)?

Several further comments are called for at this point. As I have emphasized, the criterion of p -value less than 0.05 that is often accepted in medical journals is far from adequate for such an extraordinary claim. Byrd, in order to get published in a scientific journal, was forced to acknowledge that his result is not significant when all variables are considered. He claims to improve this with a "multivariant" analysis (the correct term is "multivariate"). However, decades of experience in experimental physics have taught me that no complex statistical analysis will give convincing results when those results are not already significant based on a simpler, straightforward analysis. The only honest way to improve poor statistics is with more data.

A detailed critique of the Byrd experiment has been given by internist Gary Posner. He writes, "The most striking flaw in this study's methodology is one also forthrightly acknowledged by Byrd. It was assumed that some of the patients in both groups would be prayed for by the people not associated with the study; this was not controlled for."⁵¹

I am not so concerned as Posner about this point. It is true that apologists can use this to "explain" why the experiment did not give strong evidence for the efficacy of prayer, and this makes the prayer hypothesis nonfalsifiable and thus not very scientifically credible. However, suppose that the prayed-for group had improved so significantly that the chance explanation had been very unlikely, say with a p -value of 0.01 percent. Then we would have good evidence that the prayers of born-again Christians are not only effective but far more effective than your run-of-the mill prayer, say by a Methodist or a Catholic. The experiment did not produce this profound result,

of course, but its hypothetical possibility at least demonstrates that a scientific study of the efficacy of prayer can be, in principle, definitive.

A more serious flaw in the methodology of the Byrd experiment has been uncovered by Irwin and Jack Tessman. They discovered that the experiment was not blind after all. The code used to keep the experimental group indistinguishable from the control group was broken when Byrd wrote the first manuscript of his paper. In order to answer criticisms and produce a new version of his manuscript, he had to decide again which group did better, and this was now done unblinded. Furthermore, it was revealed that the coordinator of the study, who kept the detailed records on the patients, was completely unblinded all along.⁵²

The media hype surrounding Byrd's results undoubtedly led millions to think that science had demonstrated that prayer works, and, by implication, that God exists. However, these conclusions can hardly be drawn from this abysmal piece of work.

In 1999, another study along the same line as Byrd's was published in a major medical journal, *Archives of Internal Medicine*, by W. S. Harris and nine coauthors.⁵³ They report that the length of stay in the coronary care unit decreased by 9 percent in the prayer group, but with very low statistical significance. The p -value for their result, according to their own calculation, was over 0.25 . The total length of stay in the hospital actually increased by nine percent for the prayer group, but this was also insignificant, with a p -value of 0.5 .

Using a standard method for scoring cardiac treatment outcomes based on thirty-four adverse conditions, the investigators obtained an 11 percent advantage for the prayer group with a p -value of 0.04 . Although the paper was published, the authors admitted that their result could be simply chance.

When Harris and his collaborators applied the Byrd criteria, they obtained insignificant differences, with p -value 0.33 compared to Byrd's reported 0.05. If you were to take these statistical estimates seriously, you would have to conclude that Harris and his collaborators have disconfirmed Byrd. A more reasonable conclusion is that neither experiment showed anything significant, one way or the other.

In an amusing postscript to these studies, psychologist Nicholas Humphrey has looked at the Harris results and discovered a far more significant effect in the data than

that reported by the authors. The original sample contained 1,013 patients, and these were all randomized to determine who would and who would not be prayed for. However, the praying took a day to organize, so the researchers removed from their sample those patients who were released in less than twenty-four hours, leaving 990 in the two care groups. Humphrey found that 18 of the 484 patients who were going to be prayed for were released--before the praying started--while only 5 of the 529 who were not going to be prayed for left the hospital in that first day. The *p*-value for this difference, according to Humphrey, is 0.1 percent. He concludes that "either the study has come up with strong evidence of prayer producing *backward causation* of recovery, or else there was something wrong somewhere with the way the study was conducted."⁵⁴

Actually, they were never prayed for, so it can't be attributed to backward causation, which Humphrey obviously proposed tongue-in-cheek. I conclude that we have yet another example of how statistics can fool you when you analyze data in all sorts of combinations until you find something that looks interesting. Once again we see the filedrawer effect in action.

After the above was written and submitted to the publisher, another study on the effect of intercessory prayer on coronary patients was reported. The abstract reads:

The results of 26 weeks of intercessory prayer, a widely practiced complementary therapy, were studied in 799 patients randomized to an intercessory prayer group or to a control group after discharge from a coronary care unit. As delivered in this study, intercessory prayer had no significant effect on specifically defined medical outcomes, regardless of risk status.⁵⁵

The Tragic Story of Elisabeth Targ

Clearly we must be highly skeptical of any scientific claim that is made by "true believers," whether the beliefs being tested are religious or secular. Even the most sincere investigators may unconsciously select data that support their beliefs and ignore the data that do not, when those beliefs are deeply held. Yet another, very tragic, story

needs to be mentioned to emphasize this point.

In 1998, psychiatrist Elisabeth Targ and her collaborators published a paper in *Western Journal of Medicine* which claimed that various forms of "distant healing," including prayer and "psychic healing," significantly improved the health of patients with advanced AIDS. They reported on what they claimed was a "double blind randomized trial" involving 40 patients in the San Francisco area. Here is how the authors summarized their results, with *p*-values (*P* =) indicated:

At six months, blind medical chart review found treatment subjects acquired significantly fewer new AIDS defining illnesses (0.1 vs. 0.6 per patient, *P* = 0.04), lower illness severity (severity score 0.8 vs. 2.65, *P* = 0.03), required significantly fewer doctor visits (9.2 vs. 13.0, *P* = 0.01), fewer hospitalizations (0.15 vs. 0.6, *P* = 0.04) and fewer days of hospitalization (0.5 vs. 3.4, *P* = 0.04). Treated subjects also showed significantly improved mood compared to controls (change in POMS [a measure of mood] -26 vs. +14, *P* = 0.02). There were no significant differences in CD4+ counts.⁵⁶

The fascinating tale surrounding this experiment and its aftermath was related in a recent article by Po Bronson in *Wired* magazine.⁵⁷ Targ was the daughter of famous parapsychologist Russell Targ who conducted experiments on extrasensory perception in the 1970s that were published in the journal *Nature*.⁵⁸ In my earlier book, *Physics and Psychics*, I discussed these experiments and how they were successfully refuted.⁵⁹ Russell is a true believer and Elisabeth grew up with a firm conviction that the mind possesses paranormal powers.⁶⁰ Indeed, Russell made much of her apparent psychic powers in his 1985 book *The Mind Race*, written with Keith Harary.⁶¹ While obtaining a degree in conventional psychiatry, the paranormal remained of interest to Elisabeth and she found willing support and collaboration from parapsychologists to conduct studies in distant healing.

As a direct result of her promising 1998 study, published in a peer-reviewed journal, Targ received \$1.5 million in grants from the U.S. National Institutes of Health's

Center for Complementary and Alternative Medicine. The grants provided support for continuing distant healing studies on AIDS and further work on the brain cancer Glioblastoma multiforme (GBM). While rare, GBM is one of the most malignant forms of cancer with survival rates of about a year even with the most drastic medical intervention. Patients were dying and it seemed that anything was worth a try.

By an extraordinary coincidence, in early 2002 Targ was diagnosed to suffer from GBM. Surgery was unable to excise all the cancer and, understanding full well her slim chances of recovery by any medical means, Targ was unsure she wanted chemotherapy.

As word of Targ's illness spread, healing groups worldwide began to pray for her. As Bronson describes it,

Her bedroom turned into a circus. Healers from everywhere showed up wanting to help. It was rarely peaceful and quiet. There was Phillip Scott, a Lakota sun dancer who burned sage; Nicolai Levashov, a Russian psychic who waved his hands; Harriet Bienfield, an acupuncturist with rare Chinese herbs; Desda Zuckerman, an energy worker who used techniques inspired by the ancient methods of the Miwok peoples. The reverend Rosalyn Bruyere phoned often, trying to get on Targ's schedule. And, of course, there was her father, Russell, urging her to meditate, calm her mind, go to that place.⁶²

Sadly, neither science nor spirit was able to save Elisabeth Targ.

This tragic tale does not end with Targ's untimely death at 40. Bronson reports that the AIDS study published in *Western Journal of Medicine* had not been blinded—a fact not known to the journal. He says he has confirmed the following from one of Targ's coauthors, biostatistician Dan Moore, and physicist Mark Comings, who married Targ shortly before her death. Although not mentioned in the publication, the original study was designed to look at mortality rates. When Moore, broke the code he found that only one subject had died and so the mortality data were meaningless. Targ and the other collaborators insisted that Moore examine the data further and look at certain

HIV physical symptoms and quality of life. He found that the treatment group did no better than the control group. In fact, in some cases they seemed to do worse. Targ urged him to keep looking. Finally, after more data mining, Moore found that the treatment group had "statistically significant" fewer hospital stays and doctor visits. Targ was at a conference at the time and excitedly announced the results.

However, as Bronson put it,

This isn't what science means by double-blind. The data may all be legitimate, but it's not good form. Statisticians call this the sharpshooter's fallacy—spraying bullets randomly, then drawing a target circle around a cluster. When Targ and Sicher wrote the paper that made her famous, they let the reader assume that all along their study had been designed to measure the 23 AIDS-related illnesses—even though they're careful never to say so. They never mentioned that this was the last in a long list of endpoints they looked at, or that it was data collected after an unblinding.⁶³

If Bronson's report is correct, the authors misrepresented their experiment, claiming they had done a blinded study and implying that their criteria were pre-selected, when, in fact, they searched many post-selected criteria until they found what they wanted to see.

The editors and referees of the journals involved in the studies I have reported have done a great disservice to both science and society by allowing these highly flawed papers to be published. This has given a false scientific credibility to the assertion that prayer or other spiritual techniques work miracles, and several best selling books have appeared that exploit that theme. Telling people what they want to hear, these authors have made millions.

Healing Words

There is yet another angle in the story of prayer and healing. In his popular book *Healing Words*, physician Larry Dossey tells about the search he conducted in which he

found "an enormous body of evidence: over one hundred experiments exhibiting the criteria of 'good science,' many conducted under stringent laboratory conditions, over half of which showed that prayer brings about significant changes in a variety of living beings."⁶⁴ One wonders why he would even count those that were not conducted under stringent laboratory conditions. In any case, "many" purportedly were.

Dossey uses as his main data source a survey by Daniel J. Benor, M.D. published in the journal *Complementary Healing Research*.⁶⁵ Dossey describes and summarizes Benor's results as follows:

Benor defined healing as the intentional influence of one or more people upon another living system without utilizing known physical means of intervention. His findings:

- Researchers have performed 131 controlled trials.
- Fifty-six of these show statistically significant results at a probability level of < 0.01 or better (that is, the likelihood that the results were due to chance was less than 1 in 100).
- Twenty-one studies demonstrate results at a probability level of 0.02 to 0.05 or better (that is, the likelihood that the results were due to chance was between 2 and 5 chances in 100).
- These experiments deal with healing effects on enzymes, cells, yeasts, bacteria, plants, animals, and human beings.⁶⁶

Dossey informs us that "ten of the studies are unpublished doctoral dissertations, two are masters' theses, and the rest are published primarily in parapsychological journals." He claims that "these publications have peer review standards as rigorous as many medical journals." Note that Dossey, as quoted above, also misinterprets the meaning of *p*-values.

I have already made clear what I think of the standards of medical journals, allowing that these can be somewhat low so as to make sure that useful therapies are not kept from needy patients for too long. Unlike physicians, parapsychologists are not in the business of saving lives but rather that of investigating extraordinary

phenomena. In *Physics and Psychics* I argued that those who search for psi phenomena should be bound by the stricter standards of physics, which also deal with extraordinary phenomena. If these standards were applied, none of the 131 trials mentioned above would be publishable.

Dossey is certainly wrong when he interprets the above as "simply overwhelming [evidence] that prayer functions at a distance to change physical processes in a variety of organisms, from bacteria to humans."⁶⁷ Even without examining the detailed protocols of these experiments, the statistical significance is simply insufficient to draw such a conclusion. We have no idea how many experiments may have been done that gave no positive effects and consequently were never published. As with the Bible code, do enough trials and you will eventually get to see what you want to see.

Dossey acknowledges the problems of dealing with human subjects, which can make the interpretation of experimental data very difficult. How can one get a pure sample of people who no one prays for? The pope prays for everybody, even atheists. My Christian friends pray for me, both for my health and in the hope that I will eventually "see the light." Dossey thinks that the answer is to focus on nonhuman subjects: bacteria, mice, and even cells. However, even in this case, as Hector Avalos points out, there are people praying for the well-being of all life on Earth.⁶⁸

As we have seen, because of the difficulty of obtaining a pure, prayer-free sample, the prayer hypothesis does not have the feature of falsifiability that lends considerable credibility to any scientific hypothesis. However, not all scientific hypotheses are falsifiable, so it would be unfair for me to rule out the prayer hypothesis on that basis. What this means, in practice, is that the type of negative results we have seen so far cannot be used to *prove* that prayer does not work. On the other hand, convincing positive results with humans or other lifeforms could have, in principle, provided a good indication for some kind of nonnatural force in action. This has not happened.

Notes

1. Peter W. Stoner, *Science Speaks: An Evaluation of Certain Christian Evidences*, (Wheaton, Ill.: Van Kampen Press, 1952); Josh McDowell, *Evidence That Demands a Verdict* (San Bernardino, Calif.: Here's Life Publishers, 1972); Douglas Geivett and Gary Habermas, eds., *In Defense of Miracles: A Comprehensive Case for God's Action in History* (Downers Grove, Ill.: InterVarsity Press, 1997).
2. The fact that the neutrinos from the Sun have a lower intensity than predicted by the standard solar model is a long-standing puzzle. It now appears that this results from a phenomenon called "neutrino oscillation," a subject I personally worked on for many years and played a small role in its eventual demonstration. This serves as a good example of how the failure of a prediction in physics helps us gain new knowledge as we search for explanations of the failure.
3. For my more detailed discussion of this example, see Victor J. Stenger, *Physics and Psychics: The Search for a World Beyond the Senses* (Amherst, N.Y.: Prometheus Books, 1990).
4. "False Prophecies, Broken Promises, and Misquotes," [online], www.skepticsannotatedbible.com/prophecy.html. Thanks to Bob Phillipoff and Farrell Till for pointing out these particularly striking examples.
5. William Harwood, *Mythology's Last Gods* (Amherst, N.Y., Prometheus Books, 1992), pp. 169-70.
6. Doron Witztum, Eliyahu Rips, and Yoav Rosenberg, "Equidistant Letter Sequences in the Book of Genesis," *Statistical Science* 9 (1994): 429-38.
7. Maya Bar-Hillel, Dror Bar-Natan, and Brendan McKay, "The Torah Codes: Puzzle and Solution," *Chance* 11, no. 2 (1998): 13-19.

8. Robert E. Kass, editorial comment on Witztum et al, *Statistical Science* 9, no. 3 (1994): 306.
9. Michael Drosnin, *The Bible Code* (New York: Simon and Schuster, 1997).
10. Gans, "Public Statement," [online], members.nbci.com/bcodes/public2.htm.
11. Rips, Eliyahu, "Public Statement," [online], members.nbci.com/_XOOM/bcodes/public.htm.
12. Brendan McKay, Dror Bar-Natan, Maya Bar-Hillel, and Gil Kalai, "Solving the Bible Code Puzzle," *Statistical Science* 14, no. 2 (1999): 150-73.
13. Robert E.Kass, "Introduction to 'Solving the Bible Code Puzzle' by Brendan McKay, Dror Bar-Natan, Maya Bar-Hillel, and Gil Kalai," *Statistical Science* 14 no. 2 (1999): 149.
14. McKay, et al., "Solving the Bible Code Puzzle."
15. A Web search here will lead the reader to many sites promoting all sides of this issue (there are more than two) and provide considerable insight into personal motivations.
16. Brendan McKay, Web site [online], cs.anu.edu.au/people/bdm.
17. David E. Thomas, "Hidden Messages and the Bible Code," *Skeptical Inquirer* 21 no.6 (1997):30-36.
18. Allyn Jackson, review of *The Bible Code*, by Michael Drosnin, *Notices of the American Mathematical Society* 44, no.8 (1997): 935-38.

19. Shlomo Sternberg, "Comments on *The Bible Code*," by Michael Drosnin, *Notices of the American Mathematical Society* 44, no. 8 (1997): 938-39.
20. Drosnin, *The Bible Code*, p. 38.
21. *Ibid.*, p. 28.
22. *Ibid.*, p. 53.
23. Bar-Hillel, et al., "The Torah Codes: Puzzle and Solution."
24. Jackson, review of *The Bible Code*; Drosnin's predictions taken from *The Bible Code*.
25. Patrick Glynn, *God: The Evidence* (Rocklin, Calif.: Prima Publishing, 1997).
26. James Randi, *The Faith Healers* (Amherst, N.Y.: Prometheus Books, 1987).
27. Seth M. Asser and Rita Swan, "Child Fatalities from Religion-motivated Medical Neglect," *Pediatrics* 101, no. 4 (1998): 625-29.
28. Dale A. Matthews, Michael E. McCullough, David B. Larson, Harold G. Koenig, James P. Swyers, and Mary Greenwold Milano, "Religious Commitment and Health Status: A Review of the Research and Implications for Family Medicine," *Archives of Family Medicine* 7, no. 2 (1998): 118-24.
29. Richard P. Sloan, "Religion, Spirituality & Medicine," *Freethought Today* 17, no. 2 (2000):10-12.
30. Francis Galton, "Statistical Inquiries into the Efficacy of Prayer," *Fortnightly Review*

- 12 (1872): 125-35; [online], www.abelard.org/galton/galton.htm.
31. Physicist and Anglican priest John Polkinghorne has suggested that royalty lived shorter lives in the nineteenth century because they were more exposed to the medical care of the time.
 32. Matthews, et al., "Religious Commitment and Health Status."
 33. Harold G. Koenig, *The Healing Power of Faith: Science Explores Medicine's Last Great Frontier* (New York: Simon and Schuster, 1999), p. 24.
 34. R.P. Sloan, E. Bagiella, and T. Powell, "Religion, Spirituality, and Medicine," *Lancet* 353, no. 9153 (1999): 664-67.
 35. Linda Gundersen, "Faith and Healing," *Annals of Internal Medicine* 132, no. 2 (2000): 169-72.
 36. G.W. Comstock and K. B. Partidge, "Church Attendance and Health," *Journal of Chronic Disease* 225 (1972): 665-72.
 37. G. W. Comstock and J. A. Tonascia, "Education and Mortality in Washington County, Maryland," *Journal of Health and Social Behavior* 18 (1977): 54-61.
 38. Deborah D. Danner, David A. Snowdon, and Wallace V. Friesen, "Positive Emotions in Early Life and Longevity: Findings from the Nun Study," *Journal of Personality and Social Psychology* 80, no. 5 (2001): 804-13.
 39. Kevin Courcay, "Religiosity and Health," *Scientific Review of Alternative Medicine*. 3, no. 2 (1999): 70-76.
 40. Harold G.Koenig, Linda K. George, Keith G. Meador, Dan G. Blazer, and Peter

- Dyck, "Religious Affiliation and Psychiatric Disorder among Protestant Baby Boomers," *Hospital and Community Psychiatry* 45, no. 6 (1994): 586-96.
41. Harold G. Koenig, et al., "Religious Affiliation and Major Depression," *Hospital and Community Psychiatry* 43, no. 12 (1992): 1204.
 42. R. W. Hood, "The Usefulness of the Indiscriminately Pro and Anti Categories of Religious Orientation," *Journal for the Scientific Study of Religion* 17 (1978): 419-31; C. E. Ross, "Religion and Psychological Distress," *Journal for the Scientific Study of Religion* 29 (1990): 236-45.
 43. B. Bottoms et al., "In the Name of God: A Profile of Religion-Related Child Abuse," *Journal of Social Issues* 51, no. 2 (1995): 85-111; D. Capps, "Religion and Child Abuse: Perfect Together," *Journal for the Scientific Study of Religion* 31 (1992): 1-14; M. A. Persinger, "'I Would Kill in God's Name': Role of Sex, Weekly Church Attendance, Report of a Religious Experience, and Limbic Liability," *Perception and Motor Skills* 85, no. 1 (1997): 128-30.
 44. Kenneth I. Pargament, Harold G. Koenig, Nalini Tarakeshwar, and June Hahn, "Religious Struggle as a Predictor of Mortality among Medically Ill Elderly Patients," *Archives of Internal Medicine* 161, no.15 (2001): 1881-85.
 45. Herbert Benson, *Timeless Healing* (New York: Fireside, 1996).
 46. Irwin Tessman and Jack Tessman, review of *Timeless Healing*, *Science* 276, no. 2 (1997): 369-70.
 47. S. Walker, J. S. Tonigan, W. R. Miller, S. Corner, and L. Kahlich, "Intercessory Prayer in the Treatment of Alcohol Abuse and Dependence. A Pilot Investigation," *Alternative Therapies in Health and Medicine* 3, no. 6 (1997): 79-86.

48. Randolph C. Byrd, "Positive Therapeutic Effects of Intercessory Prayer in a Coronary Care Unit Population," *Southern Medical Journal* 81, no. 7 (1988): 826-29.
49. Ibid.
50. E. Adler, "Prayer Helps Sick in Mysterious Ways, Study Suggests," *Kansas City Star*, October 24, 1999.
51. Gary P. Posner, "God in the CCU?" *Free Inquiry* 10, no. 2 (1990): 44-45.
52. Irwin Tessman and Jack Tessman, "Efficacy of Prayer: A Critical Examination of Claims," *Skeptical Inquirer* 24, no. 2 (2000): 31-33.
53. W. S. Harris, M. Gowda, J. W. Kolb, C. P. Strychacz, J. L. Vacek, P. G. Jones, A. Forker, J. H. O'Keefe, and B. D. McCallister, "A Randomized, Controlled Trial of the Effects of Remote, Intercessory Prayer on Outcomes in Patients Admitted to the Coronary Care Unit," *Archives of Internal Medicine* 159 (1999): 2273-78.
54. Nicolas Humphrey, "The Power of Prayer," *Skeptical Inquirer* 24, no. 3 (2000): 61.
55. Jennifer M. Aviles, Sr Ellen Whelan, Debra A. Hernke, Brent A. Williams, Kathleen E. Kenny, W. Michael O'Fallon, and Stephen L. Kopecky, "Intercessory Prayer and Cardiovascular Disease Progression in a Coronary Care Unit Population: A Randomized Controlled Trial," *Mayo Clinic Proceedings* 76, no. 12 (2001): 1199-1203.
56. Fred Sicher, Elisabeth Targ, Dan Moore II, and Helene Smith, "A Randomized Double-Blind Study of the Effect of Distant Healing in a Population With advanced AIDS," *Western Journal of Medicine*, 169(6): 356-63 (1998).
57. Po Bronson, "A Prayer Before Dying," *Wired* 10.12 (December, 2002).

58. Victor J. Stenger, *Physics and Psychics: The Search for a World Beyond the Senses*. (Amherst, N.Y.:Prometheus Books ,1990), pp. 186-92.
59. Victor J. Stenger, *Physics and Psychics: The Search for a World Beyond the Senses* Amherst, N,Y,: 1990).
60. Martin Gardner, "Distant Healing and Elisabeth Targ," *Skeptical Inquirer* 25(2), 12-4 (2001).
61. Russell Targ and Keith Harary, *The Mind Race: Understanding and Using Psychic Abilities*. (New York: Balantine Books, 1985), pp. 103-16. Paperback. Original hardcover published by Villard Books, 1984.
62. Bronson, "A Prayer Before Dying,"
63. Ibid.
64. Larry Dossey, *Healing Words: The Power of Prayer and the Practice of Medicine* (San Francisco: Harper, 1993), p. xv.
65. Daniel J. Benor, "Survey of Spiritual Healing Research," *Complementary Medical Research* 4, no. 1 (1990): 9-33.
66. Dossey, *Healing Words*, p. 211.
67. Ibid., p. 2.
68. Hector Avalos, "Can Science prove that Prayer Works? The Real Story Behind the Hype," *Free Inquiry* 17, no. 3 (1997): 27-31.