photo 0560

Superstition and the Regression Effect

Whenever two variables are imperfectly correlated, an extreme value on one is likely to be matched by a less extreme value on the other. People's misunderstanding of this statistical fact results in a variety of superstitious beliefs, from the benign to the pernicious.

JUSTIN KRUGER
KENNETH SAVITSKY
THOMAS GILOVICH

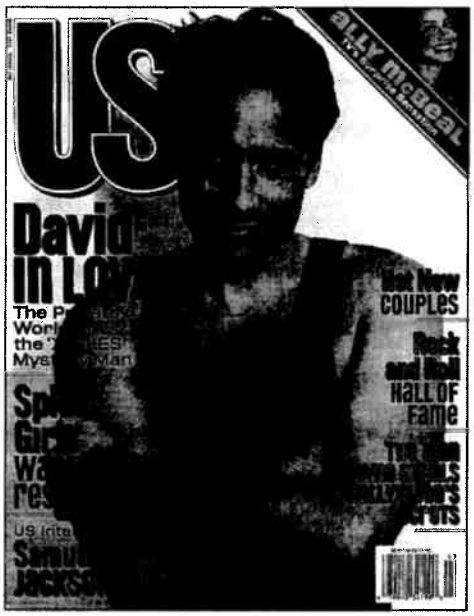
What do actors David Duchovny, Ben Affleck, and Matthew McConaughey have in common? Fame. Fortune. A penchant for science fiction films about the apocalypse. And oh yes—each has been cursed.

Consider the evidence. A recent issue of *Newsweek* noted that each actor has been featured in a different periodical in what could be mistaken for identical photographs. The shots of *X-Files* star David Duchovny (appearing in *Us* magazine) and *Armageddon*'s Ben Affleck (in *GQ*) bear an uncanny resemblance to a photo of *Contact* star Matthew McConaughey that appeared some time ago in *Vanity Fair,* when he was widely touted as the "hot" new actor on the

Hollywood scene.

And indeed, there is a resemblance: All three actors can be seen in nearly identical poses, sprawled across the front seats of old pickup trucks in jeans, rumpled shirts, and meticulously tousled hair. But it was not this eerie similarity that caught our eye. Instead, it was a warning offered by *Newsweek* to Duchovny and Affleck: "In his post-pickup phase," the article notes, McConaughey "has sometimes had a difficult time pleasing the critics" (Sigesmund 1998, 91).

The implication of the warning, presumably, is that Duchovny and Affleck should take heed of the fact that McConaughey's rise to fame became somewhat less meteoric following the publication of his pickup-truck photo. Furthermore, the insinuation is that the photograph may have had something to do with it—and that the same thing could happen to the other two.



Famous actors, athletes, and other performers are likely to experience a slump in their careers following magazine appearances like this one in *Us*.

To be sure, *Newsweek* offered its warning in jest. With a few notable exceptions, one's fame is unlikely to be affected by activities in the front seat of an automobile. But it is the case that film critics lost some of their taste for McConaughey shortly after his photo appeared in print—no longer is he Hollywood's "golden boy." And what is more, we confidently predict that a similar fate is likely to befall the other two actors. They too will likely find their celebrity fleeting. Importantly, however, we base our prediction not on any "pickup-truck curse," but on a mathematical truism known as the *regression effect.*

The Regression Effect

In the late nineteenth century, Sir Francis Galton conducted a rather uninteresting study that just happened to produce a result "of timeless significance" (Edwards 1993, 96). Galton was interested in the relationship between the heights of fathers and their sons. The uninteresting part is that he found one: Tall fathers have taller sons than do short fathers. (Of course, the same holds true for fathers and daughters, as well as for mothers and their children.) Galton's correlation was less than perfect, but it was there, just as expected.

But Galton also noticed something quite unexpected (Edwards 1993). Despite the strong relationship between the heights of fathers and sons, the tallest fathers tended to have sons that were somewhat shorter than the fathers themselves (Galton 1885). Likewise, the shortest fathers had sons who, although short, were a bit taller than they were. Why this curious attenuation? As it turns out, it *has* to be true.

To see why, it is necessary to recall that height, like many things, is multi-determined. One reason people are tall is because they have inherited the genetic predisposition to be tall, hence Galton's initial finding. But height is not purely a genetic affair, and even its genetic component is not controlled by a single gene. A person's height, after all, is determined by such disparate physical features as the height of the forehead, the length of the shinbone, and the size of each vertebra. For a person to be unusually tall or short, then, a great number of things must fall into place. A very tall person must acquire the gene from the right parent for each of these physical features, as well as receive a healthy diet, plenty of exercise, and freedom from growth-stunting pathogens. Scientists refer to the contribution of all of these hard-to-prophesize elements as "random error."

What Galton recognized was that extremely tall people tend to have random error working in their favor. That is, they are tall not simply because their parents were tall, but also because they got just the right combination of genes from their parents, were well nourished as children, led healthy lives, and so on. Of course, the children of these very tall people will, in turn, benefit from a tall gene pool, but random error is unlikely to work so well in their favor. The genetic "lottery" is unlikely to award them the favorable parental gene for as many physical features, and their childhood experiences are unlikely to be as kind. The net result is that, on average, they will be shorter than their parents. (The logic is reversed for extremely short parents.)

This is the regression effect: When two variables are imperfectly related, extreme values—high or low—on one of the variables tend to be matched by less extreme values on the other. As a consequence, very tall fathers tend to have tall children, but not as tall (on average) as they are themselves; high school valedictorians tend to do well in college, but not as well

A slip in Duchovny's or Affleck's popularity is likely, not because of any pick-up truck curse, but because of the regression effect.

Justin Kruger is a doctoral student in social psychology at Cornell University. Kenneth Savitsky is assistant professor of psychology at Williams College. Thomas Gilovich is professor of psychology at Cornell University, a Fellow of CSICOP, and the author of How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life.

(on average) as they did in high school; a company's disastrous years tend to be followed by more profitable ones, and its banner years by those that are less profitable.

What does all this have to do with Duchovny, Affleck, and McConaughey? Think of the week-to-week fluctuations in an actor's fame—exemplified by gossip columnists' relentless listings of "who's hot and who's not"—as analogous to the fluctuations in height over successive generations. To focus on those celebrities who happened to have their pictures featured in national magazines (whatever the pose) is like considering only the tallest of parents: In both cases, one has sampled from the upper tail of the distribution. The three actors in question each had their pictures featured (as is typically the case) when they were particularly newsworthy—i.e., when their careers were at a peak. And since an actor's popularity at different moments in time is imperfectly correlated (there is more than a little random error there, to be sure), one can predict by regression alone that an extraordinarily "hot" moment will be followed, on average, by a somewhat less extraordinarily hot moment, just as an extraordinarily tall parent tends to be followed, on average, by a somewhat less extraordinarily tall son or daughter.

The "*Sports Illustrated* jinx" is likely the result of regression effects. Although athletes are thought to perform poorly after appearance on the cover, Shirley Babashoff (center) went on to win five medals at the 1976 Olympics.

Thus, the *Newsweek* warning may be right—but for the wrong reason. A slip in Duchovny's or Affleck's popularity is likely, not because of any pick-up truck curse, but because of the regression effect.

Although the regression effect is easy to grasp, people often have difficulty spotting its influence in everyday life. This can result in a variety of superstitious beliefs, from the benign to the pernicious. We review a sample of these in the remainder of this article. Consider, first, three examples from the world of sports.

Regression Effects in the Sports World

Professional athletes are not known to shy away from celebrity. And yet, when it comes to having their picture featured on the cover of *Sports Illustrated* magazine, a surprising number would happily go without. Why the sudden modesty? Many athletes and fans alike believe that it is bad luck to be featured on the cover of that particular publication. Despite the prestige, they fear that it will spell doom for whatever athletic success was responsible for getting them or their team on the cover in the first place. Swimmer Shirley Babashoff, for example, reportedly balked at having her picture taken for *Sports Illustrated* before the 1976 Olympics because she was afraid she would be jinxed (Gilovich 1991). (She was eventually persuaded to pose when reminded that a cover story on Mark Spitz had not prevented him from winning seven gold medals in the previous Olympic games. Babashoff herself went on to take home five medals.)

It does not take much statistical sophistication to see how regression effects may be responsible for the so-called *"Sports Illustrated* jinx." As is the case with an actor's popularity, an athlete's performances from time to time are imperfectly correlated, resulting from a mixture of true talent, situational factors, and random error. Thus, due to regression alone, one can expect an outstanding performance to be followed, on average, by a somewhat less outstanding performance. And since athletes, like actors, tend to appear on the covers of magazines when they are at a peak, an athlete's superior performance in the weeks preceding a cover story is likely to be followed by somewhat poorer performance in the weeks after. The supernatural is invoked to explain what simple mathematics handles quite nicely.

Consider another example, this time from the world of baseball. A player is called up from the minor leagues and has a brilliant rookie year in the majors, only to slip in his second season. This scenario is so familiar that it even has a name: the "sophomore slump" (Nisbett and Ross 1980).

Is the sophomore slump real? If one examines statistics such as batting-averages, fielding errors, and runs-batted-in, it becomes clear that it is not: There is no overall tendency for a major leaguer's performance to be lower in his second season than in his first. On the other hand, there is a tendency for *certain* players to experience a mysterious "slump" in their sophomore season. Any guesses as to which? It is those who had the most exceptional rookie seasons—those whose success was probably augmented by some amount of random error, and whose performance can be expected to decline, due to regression alone, in their second year.[1]

If there is no such thing as an overall sophomore slump, why does the myth persevere? Psychologist Saul Kassin suggests it is because baseball fans are not unlike the editors of *Sports Illustrated.* Just as the magazine preferentially selects the most exceptional athletes to feature on its cover, fans tend to think about and remember those players who have had the most exceptional rookie years. One remembers the dazzling debuts of Vida Blue, Mark Fidrych, and Fernando Valenzuela, for example, but not the hum-drum inaugural seasons of other, less-celebrated rookies. Thus, belief in the sophomore slump results from a tendency of fans to spontaneously select instances from the upper tail of the distribution, and then to fail to realize that the exceptional performances of these players are likely to regress in their second seasons.

A final example from the world of sports involves a sure-fire

cure for a team's poor win-loss record: Fire the coach in mid-season and start afresh. (Indeed, for some teams, the firing and hiring of coaches is a sport unto itself!) And it works. Kassin found that baseball teams that fire their managers in mid-season win a greater proportion of their games after the change in leadership than before. But is it necessarily the case that a new manager *causes* the improvement? To be sure, there are undoubtedly cases in which the dismissal of an incompetent manager does have a causal effect on the team's level of play. But the same improvement can just as easily be accounted for by regression.

Once again, consider the circumstances that are likely to surround a team's decision to switch managers. Seldom, of course, would a team consider such dire action when things are going well, nor even when things are going only somewhat poorly. Firing the manager is a tactic that is reserved for exceptionally poor performance. And since a team's performance from game to game and year to year is correlated (most of the players are the same, after all)—but imperfectly so—a less abysmal record is almost sure to follow.

In science, the mantra "correlation does not imply causation" protects against mistakenly inferring cause from co-occurrence.

Superstition and Gambling

In 1992 alone, Americans wagered over 19.4 billion dollars in Nevada and Atlantic City casinos. They got only $16.3 billion of it back (Christiansen 1993). This is, of course, by design: All casino games are games of chance with the odds virtually always in favor of the "house"—and against the gambler. As a result, there are very few "winners" in casino gambling (unless, of course, you count the casinos). To make matters worse, these losses come at an especially great cost to a large proportion of gamblers who are at or below the poverty line (Borg, Mason, and Shapiro 1990).

The interesting thing about gambling is that gamblers *know* the odds are stacked against them. Well, sort of. Psychologists have long noted that people's expectancies for personal success in chance situations are often higher than the objective probabilities warrant (e.g., Langer 1975; Golin, Terrell, and Johnson 1977). For instance, gamblers are more confident in their chances of winning when they themselves roll the dice than when a croupier rolls for them, and they are more confident when given the opportunity to "practice." Craps players may realize that the objective likelihood of rolling a "natural eight" is 1 in 36, but may nevertheless "feel" as if it is higher.

One reason for this is that gamblers, even very experienced ones, subscribe to a number of superstitions about things they can do to improve their odds (Carroll 1998; Henslin 1967). Blow on the dice before you throw them to improve your luck. Concentrate on the number you want to get in order to achieve your desired roll. Throw the dice with the left hand to turn an unlucky streak around, or, odder still, get up and walk around your chair three times.

It should come as no surprise that such techniques are wholly ineffective. But it is easy to see how a perfectly rational person might come to believe otherwise. A strategy that a player adopts after a run of "poor luck" (such as throwing the dice with the left hand or blowing on the dice) will appear to be effective, not because it is, but because one's luck is unlikely to remain poor forever. Eventually, a string a poor outcomes will regress in the positive direction. Of course, the probability of a win is not directly affected by previous wins and losses, but if one's initial performance is poor enough, improvement is almost assured. After four lost bets on the roulette wheel, for example, the odds of improvement in the next four spins is 95 percent. Thus, a gambler will see his or her performance "improve" regardless of the strategy he or she adopted—whether it was concentrating on a particular number, blowing on the dice, or spitting on the dealer.

Note that this example is different from the others we have presented. In each case thus far, the various outcomes have been correlated: A father's height is related to his son's, a baseball player's performance in his rookie season is usually a good indicator of his performance in his second year, and a celebrity's fame at one time is correlated with his or her fame at another. In gambling, however, each roll of the dice or flip of the coin is an independent event, unaffected by the previous outcome. As a result, one can expect the amount of regression, and the misunderstandings that go along with it, to be even greater.[2]

Everyday Beliefs

Lest one think such erroneous beliefs are unique to desperate gamblers or thoughtless sports fans, consider a more common superstition, one that rings true even for, well, us. Many people subscribe to the belief that calling attention to one's successes (even privately) is a sure way to invite disaster. Indeed, certain rituals have evolved to quell the anxiety that people feel in such situations (Ferm 1959). The pronouncement that the family vacation has thus far gone off without a hitch, the baby hasn't cried all day, or the stock market is on a roll, is likely to send many individuals scrambling for a piece of wood to knock.

How does regression contribute to this belief? By definition, such activity is likely to occur only when things are going well—the very time at which one's circumstances can be expected to regress and take a turn for the worse. Reflecting on one's good fortune does not cause the decline, of course, it merely co-occurs with it. In science, the mantra "correlation does not imply causation" protects against mistakenly inferring cause from co-occurrence. Everyday causal inference, however, even among those trained in statistics, is often far less sophisticated (Kelly 1967; Quattrone and Tversky 1984).

Or consider the popular wisdom among fans of the cinema that a sequel is rarely, if ever, as good as the original. We have

little doubt that this observation is true. Whether one's measure is artistic merit or box office revenue, sequels are seldom the equals of their predecessors. To be sure, there are exceptions. Many film critics consider *Godfather II* to be as good as the original (e.g., Pauline Kael 1982), and several James Bond films have grossed more (even after correcting for inflation) than the original *Doctor No.* But examples of "sequel regression" are far easier to come by. *Star Wars* grossed far more than its sequels, as did *The Addams Family, Batman, City Slickers, Fletch, Free Willy, Home Alone, Jaws, Jurassic Park, Saturday Night Fever, Speed,* and dozens of others. Furthermore, virtually all of these films received more critical acclaim than their respective sequels.

What accounts for this trend? Avid moviegoers are quick to provide numerous explanations. Film studios may devote fewer resources to the script of a sequel, relying on patrons to buy tickets because of the quality of the original. Furthermore, some of a film's quality can derive from its novelty and originality. To the extent that a sequel capitalizes on the same basic formula as the original, then, viewers are likely to feel as if they've "been there, done that." These and other factors doubtless play a role in the typical slide in quality from original to sequel. But notice that regression alone would produce such a trend even if sequels were given the most lavish budgets and studios were not so complacent the second time around. After all, there is bound to be some random error in which movies garner critical acclaim and box office success. And since it is typically the most successful movies—ones that have likely capitalized on this random error—that are made into sequels, some regression is inevitable.

Alternative Medicine

Perhaps the most tragic instances of misunderstood regression come from a domain all too familiar to the readers of SKEPTICAL INQUIRER: the fringes of alternative medicine. Time and time again, intelligent, rational people undergo medical treatments of unproven efficacy. Ginseng, for example, has been advertised nationally as able to improve physical performance and mental alertness, despite data to the contrary (Bahrke and Morgan 1994). Shark cartilage is consumed with alarming frequency, based on the questionable logic that, since sharks do not get cancer, consumption of their cartilage ought to provide the same resistance (Lane and Comac 1992; 1996). As it happens, sharks *do* get cancer, as Lane and Comac reluctantly admit (1996, 25), but that has not quelled a seven-fold increase in the commercial slaughter of sharks in some areas as a result of the fad (NCAHF 1993). Herbalist Hulda Regehr Clark suggests a somewhat more benign treatment (for sharks, at least): She recommends a concoction of black walnut hulls, wormwood capsules, and ground cloves as a cure for both cancer *and* AIDS, according to her self-published book *The Cure For All Cancers.*

At best, these treatments are ineffective. At worst, they are dangerous. The herbal cure-all "ma huang," for example, has caused numerous deaths, as has the recent trend of coffee enemas, said to treat cancer and other diseases by "detoxifying" the body (Kolata 1996). The National Heart, Lung and Blood Institute has documented several cases of kidney failure and death in people who have undergone "chelation therapy," the intravenous injection of the synthetic chelating agent EDTA, advertised as a treatment for such diverse ailments as heart disease, Parkinson's, Alzheimer's, and sexual impotence. And in addition to these direct harms, such "cures" wreck havoc indirectly, since individuals may abandon more proven curative techniques in favor of alternative approaches.[3]

Given this track record, why are so many people convinced that these regimens work? One answer has less to do with *which* therapy an individual pursues than *when* he or she pursues it. As Beyerstein (1997) notes, many diseases are inherently cyclical—they have their "ups and downs." It is no great leap to assume that patients who seek alternative therapies do not do so arbitrarily, but when they have "hit rock bottom." Understandably so: Desperate times call for desperate measures. But the timing of such desperate action is likely to contribute to erroneous beliefs about the effectiveness of the remedies they try. Not unlike the baseball team that fires its manager in the midst of a slump and subsequently experiences an upswing in its performance, moments when one's medical condition flares up (and one happens to appeal to an alternative therapy) are likely to be followed by moments of relative relief. Thus, just as a change in a team's management can appear to have done the trick, a bogus therapy can seem effective, even when it is not.[4]

Note that the regression effect can apply just as readily to beliefs about conventional medicine. Take, for example, the common notion that a diet low in saturated fat can reduce one's serum cholesterol level. It can—but not as much as is commonly believed (Moore 1989). This exaggerated faith in the benefits of a low-fat diet is understandable, however, in light of when people change their diet.[5] One can most easily be coaxed into lowering one's intake of cheeseburgers and french fries after a test reveals an exceptionally high cholesterol level. But such an exceptional result is likely to be due, at least in part, to random error. Thus, when one is tested again, after weeks or months on a restricted diet, one is likely to get a lower result. To be sure, the change may be bolstered by changes in one's diet, but random error is unlikely to augment one's result as much as before. Thus, regression makes the causal relationship between treatment and cure appear stronger than it is.

Concluding Remarks

"It's getting better all the time. (It couldn't get no worse.)"
—John Lennon and Paul McCartney

We have discussed a number of superstitious and misguided beliefs that arise from a misunderstanding of statistical regression. Notice, however, that individuals are not wholly wrong in their predictions. In many cases, the predictions people make are quite accurate: Athletes or celebrities may well experience a lull in their careers immediately after being pic-

tured on a magazine cover, exceptional ballplayers often do go south in their second seasons, teams that fire their managers do subsequently perform better, and one may in fact experience an abatement of one's symptoms after trying an alternative medical therapy.

Thus, the error that individuals make is not one of *prediction*, but *explanation*. Elaborate causal scenarios that appeal to all manner of mystical, pseudoscientific, or otherwise superfluous beliefs are constructed to explain what is a simple mathematical given. One may be well advised to remember that when things get better, it is often, as Lennon and McCartney note, because they couldn't get much worse.

Acknowledgments

We thank Nick Epley and Tom Keegan for their helpful comments on an earlier draft of this article, and Saul Kassin for sharing his data with us.

Notes

1. Note that regression is likely to be something of a unidirectional phenomenon in this case. Although one might expect a "sophomore surge," on average, for those players who had exceptionally poor rookie seasons, such players are seldom asked back to the majors for a second year—or even make it all the way through their first. Indeed, this same asymmetry likely exists in a host of other domains, for analogous reasons. For example, professional musicians whose debut release is a flop are unlikely to land a second record deal, and businesses that turn in catastrophic losses in a particular year may be forced to close.

2. The amount of regression to be expected is inversely proportional to the magnitude of the correlation between the two variables. High correlation, less regression; low correlation, more correlation. In the extreme case of perfect association between two variables, there is *no* regression.

3. Worse still, those who utilize alternative therapies may come to ignore symptoms that indicate a need for medical attention. Indeed, patients may even be *instructed* to ignore such symptoms. The National Council Against Health Fraud (1996) notes that complaints of symptoms such as nausea, diarrhea, weakness, numbness, and tingling following herbal "detoxification" treatments are often ignored because salespeople falsely tell consumers that such symptoms are normal, and are due to the "cleansing" of the body.

4. It is for precisely this reason that clinical trials necessitate a placebo control group. Without one, it can be difficult to separate the effectiveness of the therapy from the natural course of the disease. Alternative medicines, because they are often classified as "dietary supplements," are not required to be tested in this manner by the Food and Drug Administration.

5. Another reason derives from what psychologists call the "representativeness heuristic," the notion that "like goes with like" (Kahneman and Tversky 1972; Gilovich and Savitsky 1996). Because the fat on the side of a steak or on the bottom of a skillet looks like it could clog arteries, one assumes that it does.

References

Bahrke, M. S., and Morgan, W. P. 1994. Evaluating the ergogenic properties of ginseng. *Sports Medicine* 18, 229–248.

Beyerstein, B. L. 1997. Why bogus therapies seem to work. SKEPTICAL INQUIRER 21, September/October 29–34.

Borg, M. O., P. M. Mason, and S. L. Shapiro. 1990. An economic comparison of gambling behavior in Atlantic City and Las Vegas. *Public Finance Quarterly* 18, 291–312.

Carroll, W. 1998. *Superstitions: 10,000 You Really Need*. San Diego, Calif.: SD Books.

Christiansen, E. M. 1993. 1992 gross annual wager: Industry rebounds with 8.4% handle gain. *Gaming and Wagering Business* 14, 12–35.

Edwards, A. W. F. 1993. Galton, Karl Pearson and modern statistical theory. In M. Keynes Ed., *Sir Francis Galton, FRS: The Legacy of His Ideas* (pp. 91–107). Houndmills, Basingstoke, Hampshire: Macmillan

Ferm, V. T. A. 1959. *A Brief Dictionary of American Superstitions*. New York: The Philosophical Library, Inc.

Galton, F. 1885. Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland* 15, 246–263.

Gilovich, T. 1991. *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*. New York: Free Press.

Gilovich, T., and Savitsky, K. 1996. Like goes with like: The role of representativeness in erroneous and pseudoscientific beliefs. SKEPTICAL INQUIRER 20, March/April, 34–40.

Golin, S., T. Terrell, and B. Johnson. 1977. The illusion of control among depressed patients. *Journal of Abnormal Psychology* 88, 454–457.

Henslin, J. M. 1967. Craps and magic. *American Journal of Sociology* 73, 316–330.

Kael, P. 1982. *5001 Nights at the Movies: A Guide From A to Z*. New York: Holt, Rinehart and Winston.

Kahneman, D., and A. Tversky. 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3, 430–454.

Kahneman, D., and A. Tversky. 1973. On the psychology of prediction. *Psychological Review* 80, 237–251.

Kelly, H. H. 1967. Attribution theory in social psychology. In D. Levine Ed., *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–240). Lincoln, Neb.: University of Nebraska Press.

Kolata, G. 1996. On Fringes of Health Care, Untested Therapies Thrive. *The New York Times* June 17. p. A1, C11.

Lane, I. W., and L. Comac. 1992. *Sharks Don't Get Cancer*. New York: Avery.

Lane, I. W., and L. Comac. 1996. *Sharks Still Don't Get Cancer*. New York: Avery.

Langer, E. J. 1975. The illusion of control. *Journal of Personality and Social Psychology* 32, 311–328.

Moore, T. J. 1989. *Heart Failure: A Critical Inquiry into American Medicine and the Revolution in Heart Care*. New York: Random House, c1989.

National Council Against Health Fraud. 1993. *NCAHF Newsletter* Vol. 16, July/August. Loma Linda, Calif.: Author.

National Council Against Health Fraud. 1996. *NCAHF Newsletter* Vol. 19, March/April. Loma Linda, Calif.: Author.

Nisbett, R., and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, New Jersey: Prentice-Hall.

Quattrone, G. A., and A. Tversky, 1984. Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology* 46, 237–248.

Sigesmund, B. J. 1998. Pinups pop up in pickups. *Newsweek* March 2. Hollywood scene. □