STARS AND FASHIONS IN SCIENCE

by

Richard F. Green

TR 2001-8

Department of Mathematics and Statistics
140 Solon Campus Center
University of Minnesota Duluth
Duluth, Minnesota 55812

December, 2001

STARS AND FASHIONS IN SCIENCE

Richard F. Green

Department of Mathematics and Statistics

University of Minnesota-Duluth

Duluth, MN 55812

ABSTRACT

The idea of "stars" and "fashions" in science and other areas of academic life is quite often mentioned, usually in a rather offhand way. The words "stars" and "fashions" refer to different things, but they are connected in that "fashion" refers to popular work and "star" refers to the notable person who has helped popularize the work. The word "stars" is used to refer to people who do outstanding work and are leaders in their fields, but it may also refer to media figures like movie stars who are known for their personalities as much as their work. This paper will describe some articles and books that discusse the star system and fashions in science and in scholarship. Here I concentrate on stars, while a companion paper will concentrate on fashions. I will discuss work that treats both the upper level ("stars") and lower level ("mediocrities" or the "inept"). The ultimate questions, which I cannot claim to answer, are these: Is there a star system in science? If so, what is its nature? What are its causes? What are its consequences?

INTRODUCTION

A few years ago I was struck by a very interesting commentary piece, "A man for our season," published in *Nature* by Peter A. Lawrence and Michael Locke (1997). This commentary, which was an appreciation of the late Cambridge insect physiologist, Vincent Wigglesworth, used the terms "fashion" and "star system." Lawrence and Locke wrote: "Nowadays the spotlight cast by fashion, and a star system that highlights the contributions of the few, combine to limit adventures in ideas."

Not long after the *Nature* commentary, an article by Janny Scott (1997), "Scholars fear 'star' system may undercut their mission," appeared in the Arts & Ideas section of the *New York Times*. Scott, who used the words "celebrity," "trendiness" and "flashy," discussed the academic star system and suggested some causes: "the preoccupation of administrators with rankings" and the need of the press for stars to write about and to quote. Some disadvantages of the star system are that it may be expensive to hire and support stars and "it diverts attention from scholarship to the prestige attached to it." Scott referred to articles by Cary Nelson (1997) and David R. Shumway (1997) about the star system in departments of English and literary criticism. Nelson emphasized inequalities in salaries (he thinks stars are not common enough for their high salaries to matter much, but that salary differences among disciplines is more of a problem). Shumway compared academic stars with stars in the theater and movies and pointed out how interest in stars is personalized. The reader is not just interested in the ideas, but also wants to know about the man or woman who entertains the ideas. Nelson and most of the other people I whose work I have read, or to whom I have talked, think that stars are really very good at what they do and that the star system in not a bad thing.

Shortly after the *New York Times* article, an article appeared in the *Minneapolis Star Tribune* by Mary Jane Smetanka (1998) about stars at the Twin Cities campus of the University of Minnesota. The question there was how the University was going to get money to keep the high-priced stars that they had. The *Nature* commentary and the two newspaper articles made me decide to think seriously about the "star system" and "fashions" in science. I proposed to give a talk at the UMD Philosophy Department. (Some years earlier I had given a talk, "Is science the province of mediocrities?" based on a text from Ortega. Now I proposed to look at the other end of the ability scale.) I did some reading, I thought about the subject, and I wrote and talked to people who I thought would be interested in the subject, but I never prepared or gave the talk. Since I gave up reading about "stars" in science I have frequently mentioned the *Nature* commentary by Lawrence and Locke to friends and I have sent copies to some people. I keep a copy handy, but if I want to refer to some other work on academic stars I have to dig. The purpose of this paper is to collect some references to "stars" and "fashions" in science and to write down some parts of a story that I would like to tell. One reason that I am writing this now is that I have a sabbatical to write a book on optimal foraging theory and one of the striking things about optimal foraging theory is that it was very popular for a time, but then its popularity faded. What follows is a summary of my reading, thinking and conversations about stars and fashions in science.

In this paper I begin by sketching the rise and fall of optimal foraging theory, which might serve as an example of the operation of fashion in science. Then I will discuss the

Nature commentary by Lawrence and Locke, a paper that I recommend to anyone with an interest in the operation of science. I will also discuss briefly the papers by Nelson and by Shumway about stars in literary criticism. Serious scholarly work on social stratification in science has been done by sociologists of science, especially by Robert K. Merton and his students and colleagues. I will discuss some of the ideas about scientific stars in a book by Jonathan R. Cole and Stephen Cole (1973). I will also look at the other side of the talent spectrum by discussing the 1967 article by William J. Goode, "The protection of the inept." Finally, I will try to draw some conclusions, especially about the consequences of a star system and fashions for science.

THE RISE AND FALL OF OPTIMAL FORAGING THEORY

Optimal foraging theory was an important part of the evolution-based behavioral ecology that has been expounded in a series of volumes under the title "Behavioural Ecology," edited by Krebs and Davies (1978, 1984, 1991, 1997). Optimal foraging theory, which studies the strategies that animals use (should use!) to find and choose prey, was popular for about a decade (Pyke 1984), beginning in the mid-1970s, but the popularity began to fade in the 1980s (Krebs and Davies 1991; Gross 1994). Krebs and Davies (1991) use the word "fashion" in referring to the switch in popular interest from foraging to sex:

> Sociologists of science will no doubt write PhD theses on the fact that the fashion in behavioural ecology has swung from the energetic costs and benefits of resource acquisition in the oil-crisis-dominated 1970s to sex and disease in the aids-crisis-

dominated late 1980s. (Krebs and Davies 1991, p. ix)

They were being facetious, but there is something to their suggestion that the temporary interest in optimal foraging theory was a matter of fashion.

A scientific "fashion" is a remarkably popular subject, one which is perhaps more popular than it should be for a time and then becomes less popular. People are interested in fashionable subjects because other people are interested. This point occurred to me when I was preparing a talk about optimal foraging theory to be given at Oxford in 1983. I knew that some in my audience were interested in the subject, but I wanted to tell those who were merely acquainted with the subject why they should be interested. The best suggestion that I could come up with was that they should be interested in the subject because many people, including their friends and colleagues, were interested.

People who developed optimal forging theory liked to think in terms of "strategies" that animals might use. The earliest theorists were ecologists (MacArthur and Pianka 1966; Emlen 1966; and those whose work was reviewed by Schoener 1971) who were interested in problems such as use of resources, territory size and competition. Optimal foraging theory drew a wide audience because it made the old ecological problem of animal food habits more interesting by suggesting new ways to organize data on diets and because it made testable predictions; for example, that predators should be more selective when the density of prey, particularly density of better quality prey, is greater. It is important to note that the ecological questions that motivated the earliest work on optimal foraging theory were not the same as the behavioral questions that were asked once the theory had been developed.

Optimal foraging theory rose and fell in popularity as part of behavioral ecology as developed at Oxford and promoted by two of its prominent practitioners, John Krebs and Nick Davies, who edited a series of books containing collections of wide-ranging, well-written articles about behavioral ecology. The term "behavioral ecology" has been used to refer to three kinds of work. (1) Ethological behavioral ecology (in the sense of Hinde 1982) refers to behavioral work in which differences in behavior between species or between populations is understood in terms of differences in ecological context. (2) At the other extreme, there is ecological behavioral ecology (as sketched by MacArthur 1972, or as illustrated in the symposium of the British Ecological Society edited by Sibly and Smith 1985), which attempts to use an understanding of individual behavior in order to understand its population consequences. (3) Between ethological (behavioral) behavioral ecology and ecological behavioral ecology lies behavioral ecology in the sense of Krebs and Davies, which might be termed evolutionary behavioral ecology. Optimal foraging theory became part of this version of behavioral ecology, which quickly became the predominant version and now is what most people have in mind when they think of behavioral ecology. For the sake of this discussion I will refer to the first two versions of behavioral ecology as ethology and population ecology and reserve the term behavioral ecology for the sense of Krebs and Davies.

Evolutionary behavioral ecology drew much of its inspiration from David Lack, whose work on the optimal clutch size in birds served as a model for the study of animal behavior in terms of strategies. In its early days, behavioral ecology was closer to ethology than to population ecology, but it differed from ethology in its subject matter and its method. Ethologists were more interested in behavior itself, while behavioral ecologists were more interested in the *results* of behavior. For example, ethologists were interested in courtship behavior (what behavior animals used to win a mate), while behavioral ecologists were interested in mate choice. Roughly speaking, ethologists were interested in mechanisms of innate social behavior, studied under natural, or semi-natural conditions. Ethologists were interested in functional (evolutionary) explanations of behavior to some extent, but they were more concerned with showing that the behavior that they observed was adaptive than they were in considering a variety of behavior that might be used in a particular situation and determining which was best.

Behavioral ecology distinguished itself from ethology by considering different problems, including foraging behavior, and by introducing different ways of thinking about them, including the use of mathematical models and optimization. Much of the success of behavioral ecology (both in the work that it produced and in its popularity) was due to the new problems that it considered and the new methods that it promoted. In particular, behavioral ecologists treated functional explanations much more seriously than had ethologists. However, in separating itself from ethology on the one hand and from ecology on the other, behavioral ecology lost some breadth that might have made it more stable. For example, in their important review of foraging theory, Stephens and Krebs (1986) do not mention the ethological approach to foraging (for example, they do not cite Curio 1976), and they mention the ecological consequences of the theory only to dismiss them because the theory that might be used to study them had not been sufficiently well confirmed. In my paper, "Putting ecology back into optimal foraging theory" (Green 1990), I try to argue that optimal foraging theory would have been better if it had paid more attention to the problems set by nature and if it had continued beyond finding the optimal strategies to investigate the population consequences of using these strategies.

Optimal foraging theory was begun with the hope that it would help answer questions in population ecology. For example, if animals are less selective in their diets when food is scarce, as the theory suggests, and if competition between species can be measured in terms of overlap of their diets, then species might be expected to compete more strongly when food is scarcer (MacArthur 1972). The two basic foraging models most often studied were described in two papers by Charnov (1976a, b), and I trace the subject to him, although in his excellent personal history of the subject, Schoener (1987) pointed out that Charnov's model for prey choice had been described earlier by several people. The review by Pyke, Pulliam and Charnov (1977) listed four problems that a forager faces: (1) what kinds of food to eat; (2) what kinds of patches to search; (3) how long to stay in each patch; and (4) how to move. Most of the work on optimal foraging theory was done on problems (1) and (3)—Charnov's two problems—and these are the two problems that Stephens and Krebs (1986) offer as the substance of optimal foraging theory, although Stephens and Krebs combine problems (1) and (2) into one, which they call the "prey problem." They call problem (3) the "patch problem."

By the time that Stephens and Krebs (1986) wrote their review, interest in optimal foraging theory was probably declining. Almost all of the theoretical work had been done by the early 1980s. The experimental work was continuing, but only by testing predictions from the earlier, basic models. While the subject was interesting to those who were working on it, it was not so interesting to outsiders. Certainly there is not much in the ideas tested so often in the work cited by Stephens and Krebs: Do foragers prefer the best prey? Are foragers more selective when there is more prey? Do foragers stay longer in better patches? Do foragers stay longer in all patches when travel time to the next patch is longer? Yes, yes, yes and yes, but so what? (Optimal) foraging theory as described by Stephens and Krebs was both narrow and shallow. It ignored the context and consequences of foraging behavior, and the theory tested was not interesting. It is not surprising that interest in the subject declined. Not only did interest in optimal foraging theory decline, but it was replaced by an interest in the complementary problem, predator avoidance (see Lima and Dill 1990), and a new approach that made it possible to consider several problems simultaneously—stochastic dynamic programming (McNamara and Houston 1986; Mangel and Clark 1986, 1988). Now it was possible to treat the problems of searching for prey and avoiding predators simultaneously. However, the new approach was not used to find the best way to find prey or to avoid predators; what it did was show how to balance the costs and benefits of a given form of foraging with the costs and benefits of a given form of predator avoidance.

The rise of stochastic dynamic programming and its promise of solving a variety of theoretical problems marked the end of most interest in optimal foraging theory. Work on foraging behavior continues at about as rapid a pace as during the heyday of optimal foraging theory, but the work is now more likely to be done by people interested in organisms. There is little interest in optimal foraging theory for its own sake. Most modern workers study foraging because it is something important that their organisms do, not because foraging is the subject and their organisms merely provide another example of how animals forage. For example, I think that the best recent example of the treatment of foraging is contained in Charles Godfray's (1994) book, "Parasitoids." This book combines a treatment of both theoretical and experimental work, but the point is to understand the foraging behavior of parasitoids, not to use parasitoids to illustrate work on foraging.

Optimal foraging theory is an example of a fashion in science. For a time it was more popular than it should have been (that is, more popular than its intellectual content justified), and then became less popular. It was not developed as fully as it might have been because its promoters were more interested in the theory itself than they were in the problems that the theory might have been used to solve. Fashions may produce misallocation of resources by concentrating too much effort on some areas while others are neglected. After a fashion fades, good work on the formerly fashionable subject may be dismissed because it is felt by those who follow fashions from a distance that the subject was discredited when in fact it was for a time overrated. One long-term benefit of fashions is that someone may be attracted to a subject because it is fashionable and be struck by an idea which is not of much use while the subject is fashionable, but which becomes useful later. Even if all the work done on optimal foraging theory during its heyday were thrown out, the idea of thinking about the best way to search for and select food is valuable enough to justify the expense and effort that was expended in the earlier days.

THE "STAR SYSTEM" AND "FASHIONS" IN SCIENCE—THE LAWRENCE AND LOCKE PAPER

Peter A. Lawrence and Michael Locke (1997) refer to a "star system" and "fashions" in science in a Commentary article in *Nature*. Their article is an appreciation of the Cambridge insect physiologist, Vincent Wigglesworth, who published nine books and 264 articles during a career of nearly seventy years. Lawrence and Locke admire Wigglesworth and the way that he did science, and they contrast his approach with the approach that seems required by the modern star system and emphasis on fashions in science. I think that anyone who is interested in the ideas of a star system and fashions in science should read the Lawrence and Locke paper for themselves, but I want to discuss their ideas and quote them at some length to help fix their ideas.

Wigglesworth was interested in solving scientific problems. He would make observations to answer questions, and use the observations to suggest new questions. Lawrence and Locke (1997) write:

> Many of his discoveries came, not from the narrow preoccupations of fashion, but from keen observations of experimental material, a much more unpredictable and therefore broader source. Nowadays the spotlight cast by fashion, and a star system that highlights the contributions of the few, combine to limit adventures in ideas.

If science consists of scientists asking questions of nature, finding or thinking of answers, and telling the story that nature has to tell, then some scientists might spend more time and effort extracting its story from nature, while others spend more time thinking about what story their fellows want to be told.

Wigglesworth knew that even published work was often wrong or misleading and that unpublished work would be even more unreliable. Therefore he would not talk about his work in progress or show interest in the unpublished work of others. Lawrence and Locke (1997, p. 757) point out:

> By contrast, modern science thrives on rumour; the stars in each field travel the world, talk at conferences, and publish and republish in the secondary literature. The result is a homogenization of opinion that can stifle originality; a reduction of the tendency of independent schools of scientist, particularly in different countries, to evolve their own approaches to problems.

The stars and trend-setters may be better informed than most of their colleagues, but their ideas are still a small subset of all the ideas that are possible.

The rumors that traveling scientists exchange may also be false. When I submitted my first paper on optimal foraging one referee said that the basis for the theory—that the long-term average rate of finding prey distributed in patches could be calculated using E(G)/E(T), where E(G) is the expected number of prey found in a patch and E(T) is the average time spent searching in a patch plus the time spent traveling between patches—was wrong. Later a friend mentioned that when he tried to teach about optimal foraging theory, one student was unwilling to pay attention because he had been told that the theory was wrong. My referee and my friend's student were mistaken. The mistake made its way into the literature (Templeton and Lawlor 1981) and was soon refuted (Gilliam, Green and Pearson 1982; Turelli, Gillespie and Schoener 1982). In fact, when I first saw the expression, E(G)/E(T), I thought that it should be E(G/T) and told people (thus perhaps starting a false rumor). Only when I decided to think seriously about stochastic foraging models did I realize my mistake. Later I learned that others had made the same mistake, but most had not persisted in it as long as I did.

Exchanging rumors is a very rapid way of spreading ideas ("knowledge" if the ideas are correct). Rapid communication is necessary in a fast-moving field, but there is a limit to how fast ideas can be exchanged if each message is to be a carefully worked-out response to the previous message. My first paper on optimal foraging theory (Green 1980), was an example of a model published by Oaten in 1977. Lima performed an experimental test of a model like mine and reported the result in 1984. I published a paper with an interpretation of one of Lima's results in 1990. If one thinks about this as a conversation, we might have conversed faster if our messages had to go back and forth from earth to Alpha Centari. If successive steps are taken only every three or four years, a subject will not progress very rapidly. Progress should be more rapid than this, but there is a lower limit to how long it takes between messages. Scientists can communicate rapidly over beers, but there is no time for difficult calculations or experimental observations. My observation of important people is that they do not always have time to get details right. I may take details too seriously. I have been told that details do not matter much, but it seems that one should deal with them carefully to see whether they matter or not.

Lawrence and Locke (1997) point out that Wigglesworth was shy and worked alone. They write:

> We might now stigmatize him as a loner. Selection committees would judge him lacking in the forceful leadership, managerial skills and salesmanship they suppose to be required in our tough world of entrepreneurial science. There is a place for aggressive scientific publicists, teachers and laboratory directors, but history has shown that original and creative minds are rare, and should be looked for also among the reserved and introspective.

If scientific talent is rare, then it is doubly bad to require that scientists be successful promoters. Not only will the emphasis on promotion distort the way that science is practiced, but selecting only people who can both conduct research and promote their results (and themselves) will limit the stock from which scientists are selected.

Lawrence and Locke point out that Wigglesworth thought that science administrators "should see their prime purpose to be the facilitation and encouragement of research." They quote him: "'Anyone who has tried both knows that administration is so immeasurably easier than research that it becomes the line of least resistance and that is why research needs encouragement.'" They attribute much of the problem with science to "the new cult of management which puts administrators at the top of a hierarchy, confusing management with leadership." They also think that, since research is more difficult and more important than management, salaries of researchers should be substantially higher than those of managers. It is not likely that managers, who determine each other's salaries as well as those of the researchers, will choose to have lower salaries. I don't think that a solution is at hand, but I think that Lawrence and Locke have pointed to a problem and to part of the cause, short-term management by self-important administrators.

In a sense the Lawrence and Locke paper makes two criticisms about modern science. One is the "star system" and "fashions" in science, which first attracted my attention, and the other is the cause: universities run as businesses by managers with little or no understanding of scientific research. Lawrence, who also discussed scientific fashions in his review of the book, "Egg and ego: an almost true story of life in the biology lab," by J. M. W. Slack (Lawrence 1999), seems more interested in stars and fashions, and Locke seems more interested in the university as business (Locke 1997). Both subjects are interesting. In this paper I am more interested in discussing fashions and stars in science, but if there is a star system in science it may well be because universities are run by managers who think that they require short-term quantitative measures to do their job. I am interested in the nature of the star system, in its cause and in its consequences. Lawrence and Locke think that the star system and fashions in science are bad things. I agree. Lawrence and Locke think of stars in science as being like movie stars, who attract interest for more than just their work. The star system damages science by producing fashions that narrow the focus of interest. The same point was made recently in an essay in the Science Times section of the New York Times by James Glanz (2001), "The Web as Dictator of Scientific Fashion." Much work in modern physics is discussed on the Internet, which may concentrate attention on particular problems (or versions of problems) and distract attention from others. Other commentators on the academic star system tend to think that stars are the best people in their fields and that stars are a good thing. They tend not to think about a "star system" or "fashions."

I will give two more examples of the use of the term "stars" in the academic world. One involves stars, basically media stars, in Literary Criticism, the other involves academic stars in general and those at the University of Minnesota-Twin Cities in particular. Then I will turn to look at scientific stars in the way they are viewed by sociologists.

STARS OF LITERARY CRITICISM

Janny Scott's article in the New York Times

The word "star" was used to describe academic celebrities in a December 20, 1997 article, "Scholars fear 'star' system may undercut their mission," by Janny Scott in the New York Times. The stars that Scott writes about are well-known people whom top universities want to hire to increase their institutional prestige. In this section I will comment on Scott's article and on two articles that she mentions. While these articles, and some other work that I will discuss, mention bad effects of the star system as well as the good, the balance comes down in favor of the good. None of the other people that I will mention promote the view of Lawrence and Locke (1997) that there is a better way to do science than that which is pursued by the stars and that the star system damages science.

Top universities, and those that want to be among the top, worry about their prestige, which in now often measured by rankings. The prestige of top academic universities is based on the quality of their graduate programs and on the quality of research that their people do. There is a limited supply of top researchers, and the demand for them exceeds the supply.

Scott (1997) points out that outstanding scholars are often offered higher salaries and more perks if they will move, and these offers are sometimes used as leverage to negotiate for higher salaries from their own universities. These offers and counter-offers, and the higher status of academic stars, lead to bitterness and resentment among ordinary faculty members who are not so well treated. The system promotes disloyalty among the stars and non-stars alike. The emphasis on stars may favor trendiness over slow, tedious, sound work and the cost of hiring stars may require more fundraising. As Scott says, some critics believe that the star system "diverts attention from scholarship to the prestige attached to it." On the other hand, those people referred to as "stars" really are good; they are productive, they attract good graduate students, they improve departments and they make it easier to raise money (some of which may be used to pay their own salaries and benefits).

Scott lists a number of causes of the rise of the star system: (1) administrators' preoccupation with rankings, (2) the rise of the conference circuit, and (3) the need of the press for celebrities. [Administrators need to hire people who are known, conference organizers want to attract speakers that people will want to hear (and who can attract financial support for the conference from funding agencies), and magazines and newspapers need things to write about that will interest their readers.]

Nelson's article on academic "superstars"

Cary Nelson (1997), Professor of Liberal Arts and Science at the University of Illinois at Urbana-Champaign, wrote an article, "Superstars," about the academic star system. Nelson's article, which was about the institutional effect of the star system, points to some imperfections in the star system, but on the whole Nelson seems to think that the star system is a good thing. Their colleagues often resent the higher-paid, better treated stars, and critics of the star system contend that its high cost has impoverished university education, but Nelson argues that the critics are wrong.

Academic superstars are in great demand and they can command much higher salaries than their ordinary colleagues. Nelson points out that ordinary professors are wrong to think that they are harmed by the high salaries of their superstar colleagues because there are really very few high-paid superstars so the overall effect of their salaries is tiny. Nelson thinks that the real problem of inequality in universities is in the differences in salaries among disciplines, with assistant professors in business, engineering, law and medicine having higher salaries than full professors in the humanities. Nelson also points to differences in salaries between tenure-track and temporary faculty, and he mentions the abysmal salaries for many of the non-academic staff at universities. Nelson thinks that there is little concern among university faculty for the real sources of inequality in part because most faculty members are unwilling to assume the political position needed to argue against class injustice. It is easier, he says, to concentrate on the high salaries of a few, highly visible, superstars.

Nelson says that besides disciplinary inequalities in salaries, the only other major structural disparity in faculty salaries is due to faculty members who become full-time administrators. Nelson does not mind that at his university deans may make half again as much as high-paid faculty members, but he thinks that when deans make twice as much, as they do at the University of Indiana, that is too much. It is interesting to contrast Nelson's sympathetic view, which does not begrudge administrators their higher salaries because their jobs are "unpleasant, difficult and exhausting" (p. 40), with the unsympathetic view of Lawrence and Locke (1997), who think that administrators should have lower salaries because their jobs require less talent than does research.

I think Nelson is right to point to disciplinary inequalities. In fact, the beneficiaries of these inequalities are almost embarrassingly willing to accept their advantages and ignore the disadvantages of others, but I think Nelson is wrong in his claim that the cost of superstars is a tiny proportion of the budget. He uses Duke University as the prime example of the successful use of superstars to improve the standing of a department. Nelson points out that Duke invested millions of dollars to upgrade its graduate programs in English and American literature and literary theory. Duke hired top-rated faculty and created multi-year fellowships for graduate students. The cost of such an impressive upgrade seems to be relatively high for one department or a few related departments. A truly modest effort would be unlikely to have a large effect.

Nelson argues that real superstars are really worth their high salaries, but these people are very rare. The problem, he thinks, is "premature glorification," that is, giving young people with real, but as yet only modest, achievements high salaries and treating them as stars. Such appointments cause resentment, not only among the unsuccessful and disaffected, with whom Nelson has little sympathy, but also among those who have achievements comparable to those of the newly coined stars. Nelson claims, I think correctly, that few people will resent the appointment and recognition of the truly outstanding. The problem which Nelson does not mention is that the truly outstanding are very rare, and the demand for people good enough to raise or maintain the reputation of a department is much higher than the limited supply can satisfy.

Nelson seems to accept the universities' commitment to the star system. He has two suggestions to reduce the resentment that this system can produce. One is to avoid premature glorification of young, rising stars. The other is to discourage the "tacky and embarrassingly self-important behavior" that superstars sometimes show. He gives some examples of such behavior and tells how to treat it.

Nelson's discussion of the star system concentrates on its effect on institutions rather than on disciplines. Therefore it is mostly about the stars and the system rather than about the fashions that stars may promote. However, he does comment briefly about the careers of stars. Just as movie stars have their fans, academic stars have theirs, but academic fans use citations rather than applause to reward their heroes. Nelson writes (p. 39):

> Citation, indeed, is somewhat like academia's version of applause; many journals will not stage your production unless you cite celebrities. Some superstars adapt to changing audience tastes and remain in the public eye, remaking themselves as the need arises. Others, superstars of long standing, may come to seem like monuments to paradigms past, as the leading edges of disciplinary debate leave them behind. Academic celebrity is not a uniform or consistent or dependable intellectual category. Its only constant may be the financial rewards it often brings.

<u>Shumway's article about stars in literary criticism</u>

David R. Shumway (1997), Associate Professor of Literary and Cultural Studies in the English Department at Carnegie-Mellon University, wrote an article, "The star system in literary studies," which looks at the same subject as Nelson, but uses a different point of view. Shumway traces the rise of the star system in literary studies and explicitly compares academic stars with other kinds of stars, especially those in the theater and in movies.

In tracing the history of stardom in the theater and movies, Shumway begins by mentioning that the theater had stars in the nineteenth century, but their stardom was different than the stardom of movie stars in the twentieth century. What the early theatrical stars had in common with movie stars was the visual presentation of their personalities, but in the earlier days this presentation was only through their performances. An important characteristic of stars is that they have fans who may imagine having personal relationships with them. When movies developed, their stars could achieve much greater celebrity than theater stars of the nineteenth century, both because movies could be seen widely, and because movies developed after the rise of the popular press.

Stars are known for their personalities as well as their performances. The spread of newspapers, glossy magazines and finally television, made it easy to see the celebrities. Other celebrities, including politicians and writers, were increasingly viewed as personalities. Finally, after the rise of television, publicists manufactured celebrities not only based on their accomplishments, but also on their ability to be famous, resulting in people who are "famous for being famous." For the last half century, celebrity has pervaded our culture. Stars are celebrities, but they are more. Stars have lives that interest people. Their fans are very interested.

Shumway (1997) thinks that the star system in literary studies arose in part because of the pervasiveness of celebrity in our society. The publicity system that came to rely on celebrities in other fields was ready to welcome academic celebrities as well. All that was needed was for literary studies and other academic fields to develop celebrities who might be promoted. Literary scholarship was dominated by philology early in the twentieth century, and its practitioners viewed themselves as scientists. They regarded their personalities as irrelevant to their work. Personalities did become important as literary criticism rose to prominence, but there were many approaches and many specialties, so it was difficult to find stars whose work was broadly based. However, the rise of theory in the 1970s made it possible for individuals to apply a single method broadly enough so that their work would be of interest to a large fan base. The combination of personal approaches required in criticism and the general appeal of theory with the hunger for celebrities in the overall culture helped lead to the star system in literary studies.

Shumway writes (p. 88):

> I am not, then, arguing that movie stardom is a direct historical cause of the star system in literary studies. I am suggesting that the importance in contemporary America of celebrity in its many forms, including stardom, helps account for the rise of the academic star system.

Other events in the university system also contributed to the rise of the star system. Universities increased rapidly in numbers and size after World War II and many more Ph. Ds. were produced. Then the Ph. D. supply exceeded demand and the job market was very bad in the 1970s and 1980s. Ph. Ds. are trained to do research, and when the job market was bad, young professors were expected to do a lot of research in order to get tenure. Some succeeded, many failed, and the most successful were the stars. [I think that there are two things going on here. One is that when there is fierce competition, some will do much better than others, thus producing stars in the sense that some are relatively outstanding. The other thing is that with a crowded field there is not enough to do. Fifty thousand professors cannot all write one or two papers a year, each doing groundbreaking work on some major writer. But if the stars are also fashion leaders, it might be possible for quite a few people to apply the newly fashionable ideas to the relatively few major literary works. It is also convenient if other, completely new subjects are introduced. Shumway writes about stars "authorizing" work. Maybe what they do is not so much convince people that their particular ideas are right as make it permissible to work in a certain way on certain problems.]

The rise of the conference circuit also contributed to the rise of the star system. One particular factor was the rise of thematic conferences with no large ready-made audience. For such conferences to succeed it was important to have speakers whose presence would attract enough attendees. What was needed was to find stars who had fans. [Even if one started without (media) stars, one could invite outstanding people, who become better known because they have been invited. Now they are more likely to be invited again and eventually begin the life of academic stars.]

Shumway concludes that the star system in literary studies damages the discipline by "inhibiting the production of collectively held knowledge" (p. 86). The star system leads to resentment of both the stars and the profession that produces them. "Moreover," he writes, "the star system reflects a shift in symbolic capital from the collective judgment of the discipline to individuals, a shift that further diminishes the discipline as a source of identification, breaking it down into incommensurable camps." He goes on, "The star system depends on fans, an impoverished community focused on individuals who are not part of the community. It would be better for literary scholars, teachers, and students to stop being fans and to recognize that they can authorize knowledge without the name of a father or mother" (p. 98).

Shumway's explanation for the rise of the star system in literary studies depends on a condition limited to that field, or at least quite limited, the fact that criticism is based on a particular point of view and this point of view (maybe the critic's, maybe Marx's or Freud's) must be "authorized." It is authorized by the readers' willingness to agree with premises chosen by the critic. An argument is not accepted because of its intrinsic logic, or by the fact that it relies on experimental observations made in a clearly described way. It is accepted because it follows from some authorized premise that the reader is willing to accept for now. Shumway thinks that because of this difference between literary studies and science that star systems have not arisen in science. He writes:

> Because authority in the natural sciences is rooted in a consensus about such norms [structural and intrinsic norms established in cosmology or physics, say], the hierarchies in these fields have not developed into star systems of the sort I have described here.

There may be a need for internal authority in literary studies that is not needed in science, but many of the other factors that affect literary studies are in play in science. Scientific ideas may be shown to be right or wrong, but whether they are interesting or important is another thing. Whether a piece of work gets published is based not only on its truth, but also on whether reviewers and editors find it interesting. One of the functions of stars in science is to tell people what is interesting.

STARS AT THE U

Universities compete among each other for top quality faculty members. Universities offer faculty members at other universities more money than they are currently making if they will move. The faculty members' home universities may make counter offers in order to keep their top faculty from deserting. Large universities with large faculties, many of whom are not particularly well paid, are susceptible to raiding by richer universities. The University of Minnesota-Twin Cities is such a large university, vulnerable to raiders. A few years ago, when the University adopted an initiative to develop molecular and cell biology, other departments expressed fear that they would lose many of their top-ranked faculty if they did not receive more money to match offers made by raiders. This problem was discussed in the article, "'U' science makes some feel short-shrifted," by Mary Jane Smetanka in the Feb. 21, 1998 issue of the *Minneapolis Star Tribune*. The article used terms such as "stars," "star faculty," and "blue-chip" faculty. The basic issue for the departments and for the university as a whole was the report on graduate education by the National Research Council which periodically ranks graduate departments. Overall the U ranked 20th in graduate education among all research universities in 1995, down from 16th in 1982. The U was still in the top ten among public universities, but it had set itself the goal of being in the top five. However, in salaries the U ranked 28th out of the top 30 research universities.

The U certainly faces a difficult problem trying to compete for stars with richer universities. Perhaps the U should resign itself to the role that it may be forced to play whether it likes it or not—serving as a stopping place for promising young scholars who will stay for a time and then move on if they succeed in their work and become famous. This is not an unimportant role to play in the larger scheme of things, but it must be frustrating to those who want the U to be the Harvard of the Upper Mississippi.

[When one's ambitions outrun one's resources it is good to increase the resources or decrease the ambitions. To do neither leads to frustration. I think that it would be best if the U were run by people who understood the purposes of a university. The U should want people who can do good work. Someone should pay attention to see whether good work is, in fact, being done. Some years ago an ecology faculty member did a study of his department and the cell biology department and concluded that within rank there were no differences in salaries due to amount or quality of research (measured in terms of citations). Having served as an administrator had a big effect. This study was reported in the *Minnesota Daily*, the student newspaper, but it drew no response from the U administration. At the same time there were newspaper articles bemoaning the fact that U faculty were being raided by other universities. I thought that the U was easy pickings. They have good faculty members, but they do not pay them very well, and they do not know who is good and who is not. Another university could look at a target faculty member and be quite sure that he or she would not be well-paid and that the U would not know whether he or she was any good. Of course, once an outside offer is made the U might suddenly realize that they had a good person on their hands who is about to get away. It would be better to know before an outside offer is made whether a faculty member is good and is worth keeping. If a university does not have much money it must have something else going for it in order to keep good faculty members.]

I asked several people mentioned in the Star Tribune article what they and others meant by "star." I got three answers and they all used the word "reputation." One referred to people getting "international reputations," another referred to people who help make a "department's reputation." The point is that a star at the U is someone who has done good enough work so that he or she is known and respected by outsiders and the respect is great enough to raise the reputation of the department. I had no sense that anyone was using the word "star" in an ironic way. I was surprised by this and by the thought that many people could use the word "star" without embarrassment. The impression that I got was that everyone thought that stars were a good thing and there was nothing strange about using the term "stars." Recently, a Star Tribune editorial ("'U' stars: medical leadership points the way," January 9, 2001) used the term "stars" in a favorable way to urge the legislature to support increased funding for the medical school. The basis for the editorial was that two U biomed researchers had been listed in the top 21 spots by U. S. News. They were the stars.

I also asked our former dean, Rip Rapp, about academic stars. I asked him what he thought constituted a scientific star. He was more specific than the people in the Twin Cities and suggested that a star was someone at the level of the National Academy of Science. He thought that maybe about five times the number of people actually in the NAS are at that level, which would mean that there are around 8000 scientific "stars" in the U. S. Rip also said that he thought that we had several "stars" at UMD. His idea of a star was someone who would be mentioned at a meeting in Europe. In fact, I think that the definition that Rip had in mind was exactly what others have in mind, a star is someone whose name pops to mind in connection with a school or department. Stars in this sense are known for their work and not for their public personalities.

The term "star" is used in a favorable way by deans and department heads who want to convince others of the value of particular faculty members. What it means is that the person is good, that we want to keep him or her. The issue is political or economic, but the quality that matters directly is reputation, which indirectly is due to the quality of work. There are two distinctions here. One is between reputation and the work that the reputation stands for. This is parallel to the distinction between the star and the good scientist. The other is between the effect of the star system on departments and universities and the effect on science or scholarship itself. I am interested in the effect of the star system on science, but this may not be understandable without considering a longer causal chain. Universities want good scientists, or scientists with good reputations. The reputation is influenced by the work. However, the star system itself influences the kind of work that is done and what of it is accepted and valued. Thus, universities influence science indirectly [and unintentionally, I would say], through the star system which they are stuck with because of competition.

An approach that looks at the scientific star system more in terms of the scientists and the work they do rather than the connection between stardom and academic institutions is offered by the sociology of science. Sociologists are interested in social stratification in general and some are interested in the social stratification in science.

SOCIAL STRATIFICATION IN SCIENCE

Sociologists have studied social stratification in science at least since the 1960s. Robert K. Merton was a leader in the development of the sociology of science. Merton studied the structure and function of institutions. He was interested in a wide variety of phenomena and advocated the use of both theory and empirical methods. Merton proposed certain "norms of science" including *universalism, communism, disinterestedness,* and *organized skepticism.* There might be arguments about the extent to which each of these ideals is actually met by science as it is practiced. The question of whether there is a star system in science involves the first norm, universalism. Universalism requires that truth claims be judged on impersonal ("universal") criteria. The opposite of universalism is particularism, of which nepotism serves as the ideal case. Cole and Cole (1973) treated the issue of universalism versus particularism by looking at two sorts of information: (1) quantitative facts about the variability in production and recognition among scientists, and (2) evidence, including statistical evidence and reports, of causes of the observed variability and the resulting apparent hierarchies in science.

I think that the best empirical approach to the star system in science is through the work of sociologists of science who followed Merton, including the Cole brothers and Zuckerman (1977), and through the use of statistical information of the sort reported by Price (1963) in his beautiful presentation of the ideas of what became known as the "science of science." (See also Ziman 1984.)

There are difficulties in studying the social system of science. One difficulty is the lack of data. Workers use what is available, not because it is exactly what they want, but because it is available. The data that are used do not always seem to point to the same conclusions that the sociologists reach, perhaps because it is difficult to separate cause and effect. Statistical methods have also been misused. I will point out a few examples of mistakes that have been made.

For the most part I will follow the treatment that Cole and Cole (1973) give in their chapter 3: "Patterns of stratification in American science." The first problem is to establish the existence of hierarchies in science. Cole and Cole (p. 37) begin by asserting: "There are few disagreements that science is dominated by a relatively small elite; there are disagreements about the bases on which rewards in science are distributed." Cole and Cole point out that there is a sort of hierarchy of disciplines within science with physics ranking highest, followed by chemistry, biology and geology. Within disciplines there may be a hierarchy of sub-disciplines. For example, particle physics has higher status than solid-state physics.

Individual scientists occupy positions in a hierarchy of accomplishment. At the top are the greats of history, like Newton, Darwin and Einstein. Among living scientists, the highest status is occupied by Nobel laureates. Below these are members of the National Academy of Science in the United States or Fellows of the Royal Society in Britain. These people are clearly members of a scientific elite. Many of these people are professors in major research universities. Below these people in status are other faclulty members at major research universities, and below them are faculty members at other Ph. D.-granting universities. Many Ph. D.-holders have jobs in industry or in four-year colleges or other non-Ph. D.-granting universities. Most of these people publish little or no scientific work.

Scientists vary greatly in status. The differences in status are due largely to differences in achievement. The status system may not be fair (family status and sex have huge effects on the chance to reach the top of the hierarchy, for example), but, putting questions of fairness aside, the scientific elite are clearly people of accomplishment.

Elite scientists are not only accomplished, but they have power and influence as well. They control, or at least influence, the allocation of resources. They decide who will get money (and therefore, to a large extent, who can work and what work they can do) and whose work will be published. Some of the elite are recognized as stars, who serve two roles in science: (1) they provide examples and serve as role models, and (2) they provide authority—work of lower status scientists may be validated by referring to their work.

Two different methods have been used to study the social structure of science. One is the accumulation of anecdotes. Zuckerman (1977) provides many examples in her treatment of American Nobel laureates. For example, they tend to come from middle, or upper-middle class backgrounds, they go to good schools and do their graduate and post-doctoral work under the supervision of outstanding scientists, often Nobel prize-winners themselves. This approach risks criticism as methodologically unsound, but I think that it is the most fruitful approach to understanding social influences on recruitment. In fact, Heilbrun (1977) criticized Zuckerman for not including more personal characteristics of her subjects. The other method is to compile statistical data and try to draw conclusions from the accumulated data. This method is more respected methodologically, but the accumulation of a large amount of questionable data (data from many incomparable sources, for example) should not be preferred to anecdote just because impressive analytic methods seem to be applicable.

Here I am interested in three eponymous ideas that have been used by sociologists of science to discuss the hierarchy of achievement in science. These are "Lotka's law" (Price 1963), the "Matthew effect" (Merton 1968, 1988) and the "Ortega hypothesis" (Cole and Cole 1972).

"Lotka's law"

Lotka's law is named for the American scientist, Alfred Lotka, who observed in 1926 that the number of authors who had published exactly n papers was proportional to $1/n^2$, which Cole and Cole (1973) refer to as an "inverse square law." Lotka's law can be expressed, approximately, in an even simpler way: among authors who have published at least one paper, the proportion who have published n or more is equal to 1/n. That is, one author in ten will have published ten papers or more, one in twenty will have published twenty or more, and so on. Lotka's law specifies a frequency distribution for the number of papers that authors will publish. This (theoretical) distribution has the interesting property that the mathematical expectation of the number of papers published by an author selected at random does not exist (that is, it is infinite). [I showed that this distribution could be obtained by using a random walk model (a "random excursion," actually) (Green 1986). While the expectation does not exist, any sample would have an average, and such averages would satisfy an analogue of the law of large numbers, that is, they would be relatively stable. The result is similar to that shown for the St. Petersburg problem by Feller (1968), but the stabilizing factor involves lnN for Lotka, while it is $\log_2 N$ for St. Petersburg (where N is the sample size = the number of authors).] The tail of the distribution for Lotka's law is heavy because there is a moderately large probability that an author will have a huge number of publications. In fact, the "inverse square" law slides into an "inverse cube" law for moderate numbers of publications (Price 1963), and the infinite expectation is not really a problem. However, what Price did was to consider Lotka's law up to a fixed number of publications and then truncate. (For example, no one could have more than 100 publications.)

The truncated version of Lotka's law had the interesting property that in a discipline with N authors, half of the articles would be written by about the square root of N authors (Price 1963). That is, in a discipline with 100 published practitioners, about ten (or 10%) would have published half of the articles. In a discipline with 1000 published practitioners, about 32 (or 3.2%) would have published half of the articles. Two conclusions follow from this observation. First, most of the articles in a discipline will be written by a small proportion of the practitioners (this is true even if one makes more realistic assumptions about the shape of the distribution of productivity). Second, if one defines the elite as those who publish the most papers, and if one uses as a cutoff for the elite that number of articles such that half of all the articles are published by people with that many publications or more, then as the number of practitioners in a field increases so will the number of the elite, but the proportion that are elite will decrease. (This second conclusion depends strongly on the shape of the tail of the distribution of productivity.) [I think that Lotka's law points to something of importance, that the distribution of the number of publications per author is big tailed, but some calculations such as "Price's law" are simply numerology.]

Citations also follow a law like Lotka's. In fact, the number of times that a paper is cited may fit Lotka's law better than the number of publications per author. The tail for citations may be a bit lower than Lotka's law specifies, as is the tail for the number of publications, but there is not as strong a force limiting the number of citations as there is for publications.

I can think of two simple, not completely implausible, mechanisms that might produce Lotka's law, or the simple modification of it, $f(n) = 1/[n(n+1)]$. First, if the odds that an author will publish another paper, given that he or she has already published n papers is equal to n:1, then the probability function $f(n) = 1/[n(n+1)]$ will result. It is likely to be true that the more papers that have been published the easier it is to publish another. I can see no reason why the probability of publishing another should be exactly $n/(n+1)$, but if it were, the desired result would follow. Second, if the total number of publications grew according to a pure birth process, but with the whole process stopping at some time determined at random, with a stopping rate equal to the birth rate, then a Lotka-like law would follow. Price (1963, p. 50) suggested defining a man's "solidness" as the logarithm of his life's score of papers. If it is proper to measure differences in productivity in terms of relative differences, that is, the step from 2 to 4 publications is the same as the step from 4 to 8, then Price's suggestion has merit. A pure birth process is a random process in which given relative changes tend to occur at a constant rate. Any individual can give birth, so the overall rate of growth of a population is proportional to its size. This is like saying that the next paper might flow out of any of the papers already published.

These models produce a big-tailed distribution because they result from processes in which growth depends on the size already achieved. Something like this almost certainly happens in the cases of publications and citations. However, these models make no assumption about differences among authors or papers. In fact, authors differ in ability and ambition, and papers differ in importance. Thus, a realistic model would be a mixture or distributions produced by some mechanism(s), and mixtures tend to show more variability than the distributions that are mixed. A realistic model to produce Lotka's law would require a mixture of distributions, each with less variability than Lotka's law. Such a model would require a mixing distribution and a family of distributions to be mixed. I have not been able to work out a reasonable theory. [A Pareto mixture of Poissons works, at least in the tail, but I can't justify either the Pareto or the Poisson.]

No matter what mechanism produces the distributions of the number of papers per author, or the number of citations per paper, and no matter whether Lotka's law provides a perfect fit to the data, there clearly are large differences among authors in their productivity and among papers in their visibility. A very few authors publish many papers, a relatively few publish most papers, and a large number of workers publish very few papers, or none at all. The "Matthew effect" and the "Ortega hypothesis" refer to these differences in productivity (or perhaps the resulting differences in influence). Work on the "Matthew effect" concentrates on the elite, while the "Ortega hypothesis" asks whether the work of the mass of mediocrities is of any importance to science.

The "Matthew effect"

The "Matthew effect," which was named by Robert K. Merton (1968, 1988), is basically the principle that the rich get richer and the poor get poorer. The name is taken from the gospel of Matthew, in particular verse 25:29, "For unto everyone that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath." In fact, some scientists have more than others, but it is not clear that this happens because the rich are rewarded and the poor punished. What happens is that young people who show early promise [one of the points of Merton's (1988) paper is to decry the emphasis on precocity in science] go to better schools, receive better training and the patronage of more powerful masters than those who show less promise. At each stage the better are favored, but there is little evidence that they are favored because they have more rather than because they are better able to do what is expected of them. Zuckerman (1977) does report the comments of a number of Nobel laureates who observed that they receive more credit than their coauthors when they write joint papers. This is some evidence that the rich are favored, but the amount of favor is slight.

There is quite a difference between the family background of Nobel laureates and the background of their cohorts in America. Zuckerman (1977, pp. 63-68) points out that 82% of Nobel laureates (81.7%, actually) had fathers who were professionals or managers or proprietors (as opposed to farmers, clerical workers or skilled or unskilled workers), while the percentage in the whole U. S. population was 11% (11.2%, actually). If one were to assume that there is no difference in ability of children whose fathers had different occupations, then one might expect that the proportion of people able to do Nobel-caliber work would be the same for all occupational levels. Therefore, one would expect the percentage of Nobel-caliber work among children of lower occupational class, x, to satisfy the equation: x/88.8% = 81.7%/11.2%, which implies that x = 647.8%. Since 18.3% win Nobel prizes we can subtract 18.3% from 647.8% and get about 629%, which is the percentage of additional Nobel-caliber work that we might have expected if children from all family backgrounds had been raised to the same level as the highest group. Of course, the assumption of equal distribution of ability might be false, but if is it true, it might have been possible to have seven times as much high-quality work as actually occurred. It is also possible that only a limited amount of useful work was possible within the time given and that a large fraction of that work was done, even though the man-power supply was severely limited.

The "Ortega hypothesis"

The "Ortega hypothesis" (Cole and Cole 1972) gets its name from a passage in the book, *The Revolt of the Masses*, by Spanish philosopher Jose Ortega y Gasset (1932, pp. 122-123):

> What happens is that, enclosed within the narrow limits of his visual field, he [a type of scientist unparalleled in history] does actually succeed in discovering new facts and advancing the progress of the science which he hardly knows, and incidentally the encyclopedia of thought of which he is conscientiously ignorant. How has such a thing been possible, how is it still possible? For it is necessary to insist upon this

> extraordinary but undeniable fact: *experimental science has progressed thanks in great part to the work of men astoundingly mediocre, and even less than mediocre* [my italics]. That is to say, modern science, the root and symbol of our actual civilization, finds a place for the intellectually commonplace man and allows him to work therein with success. The reason of this lies in what is at the same time the great advantage and the gravest peril of the new science and of the civilization directed and represented by it, namely mechanisation. A fair amount of the things that have to be done in physics or in biology is mechanical work of the mind which can be done by anyone, or almost anyone. For the purpose of innumerable investigations it is possible to divide science into small sections, to enclose oneself in one of these, and to leave out of consideration all the rest. The solidity and exactitude of the methods allow of this temporary but quite real disarticulation of knowledge. The work is done under one of these methods as with a machine, and in order to obtain quite abundant results it is not even necessary to have rigorous notions of their meaning and foundations. In this way the majority of scientists help the general advance of science while shut up in the narrow cell of their laboratory, like the bee in the cell of its hive, or the turnspit in its wheel.

Ortega's point was to decry the mechanisation of science and much of modern life (or what was modern life in the 1930s). Cole and Cole (1972) used the term "Ortega hypothesis" to represent the idea that much valuable research is done by ordinary scientists (Ortega's "mediocre"). The Coles believe that little research of value is done by the majority of scientists. That is, they state the "Ortega hypothesis" in order to refute it. They offer several bits of evidence in support of their conclusion. Most scientists (Ph. Ds. in science) publish few if any papers, and the few papers that they publish are not cited many times. That this is true can be seen from the data. The Coles also argue that the truly outstanding work is less likely to rely on the work of ordinary scientists than is the work of ordinary scientists. Here I think that they make some mistakes. For example, in one study (Cole and Cole 1972) they "asked a well-known physicist to list the five most important contributions to elementary particle physics in the last 10 years" (p. 371). The five articles named were found to have cited 107 American physicists, all but one of whom were located in "the top nine physics departments in the United States, or at such distinguished laboratories as Brookhaven or the Lawrence Radiation Laboratory." The Coles take this as evidence that the top researchers cite the top researchers preferentially, but it is also possible that elementary particle physics is studied at only a few places. If almost all of the relevant work is done at a few places, it would not be surprising that the work cited was done at those few places.

In another study Cole and Cole looked at the authors cited in the best papers of 84 university physicists. A sample of 385 cited authors had received an average of 119 citations in the 1965 Science Citation Index, while the average author listed in the SCI was cited only 6 times. The difference between the averages of 6 and of 119 seems huge, but in fact, this difference might simply be an artifact of the method used to calculate the average. I do not know how the numbers of citations per author are distributed, but if they have the same (or approximately the same) Lotka distribution as the number of publications per author or the number of citations per paper, then the values of 6 and 119 are about what one would expect if there were no effect of the quality of citer on the quality of paper cited. [Let me make the argument for the number of times that a paper (and not the lifetime work of an author) is cited. To simplify the calculations, I assume that the number of citations per paper will have a truncated Pareto distribution given by density function $f(x) = 1/x^2$ for $1 < x < 1000$. This is not a proper distribution since the total probability is 0.999 rather than 1.000, but I ignore this detail. That is, I use a continuous version of Lotka's law, and I assume that no paper is cited more than 1000 times. Then the average number of citations per paper is given by $\int x\, x^{-2}\, dx = \ln 1000 = 6.9$. But if we look at a citation (or a paper that receives a random citation) we must weight the probability of the cited paper having a particular number of citations by the number of citations. That is, instead of averaging over the density $f(x) = x^{-2}$ for $1 < x < 1000$, we must use the density $f(x) = x^{-1}/6.9$ for $1 < x < 1000$ and we get $\int x\,(x^{-1}/6.9)\,dx = 999/6.9 = 145$ citations per paper. That is, the average number of citations per paper is about 7 or about 145, depending on whether one averages over papers or citations. The Coles' method resembles averaging over citations, not papers (they use authors, actually). The huge difference that they point to says nothing about whether the elite perfer to cite the elite. It only shows a confusion about how to calculate an average. This point was made earlier in an ecological context by Monte Lloyd (1966) who described "average crowding" in terms of how many other individuals each individuals share a space with, but not how many individuals are in an average space (with some spaces empty or sparsely occupied).]

The Coles do give some convincing evidence that there is some difference in the types of authors that are cited by better or worse papers (Cole and Cole 1972, Table 1, p. 370, for example), but these differences are not nearly as striking as the misleading numbers given by calculating averages two different ways. There are a number of reasons that papers are cited, the best of which is to show what influenced the citing work. Papers are also cited to acknowledge the ownership of an idea and also to advertise favorite work. Often papers are cited to provide authority for the citing work. Authors sometimes cite work to show that they are aware of the literature or to placate potential referees. Better-known authors are less likely than unknown authors to cite papers for some of these reasons. Elite authors have their own authority and have less reason to appeal to the authority of others than do relatively unknown authors. And, if authors fear that they must cite their betters, elite authors have fewer betters to fear.

One issue that Cole and Cole (1973) mention is the willingness of the non-elite to accept the status of the elite. Among the reasons suggested for this acquiescence is that many members of the non-elite are socialized into respectful attitudes by their elite mentors. I think that it is possible that those slightly below the highest level may be hopeful of achieving higher status, while those at a much lower level have given up hope. The educational process for research scientists resembles a sequence of competitive games and after many games each person knows his or her standing quite well. The hierarchy of status, which is based largely on achievement, is quite natural in an area in which ability varies widely and achievement is quite easily recognized. But the system of science also requires authority and this authority is supplied by tradition, and by leaders who embody the tradition. Cole and Cole (1973, p. 79) repeat a story told by Michael Polanyi who proposed an unorthodox theory as a young chemist. His theory was rejected but it later turned out to be correct. Polanyi believed that it was right for his idea to be resisted because as a youngster he had not earned the right to have ideas so different from the prevailing orthodoxy.

I think that the view that Polanyi was supporting is seen in the traditional form of the "null hypothesis" tested in science. If there are two types of error, one more costly than the other, the more costly type is called "Type I" error. Then the "null hypothesis" is the safe hypothesis, in the sense that one can avoid the more costly error by accepting the null hypothesis. Therefore, Type I error—the worse type or error—is to falsely reject the null hypothesis. In natural and social science the null hypothesis usually has the form: there is no difference between treatments, or the traditional idea is correct. What this implies is that the worse type of error is to say something new when the old idea is true, or to say that there is a treatment effect when, in fact, the treatment makes no difference.

Talking about a star system focuses attention on the top, but what goes on at the bottom is also important. Cole and Cole (1973, p. 73) contrast the way that scientists find their place in the hierarchy with what happens in business:

> It has been said that in business world, people "rise to their level of incompetence." In science, people sink to their level of competence.

Their idea is that many scientists are trained at very good research universities. Some obtain post-doctoral fellowships or faculty positions at such top places, but most are not good enough to stay at that level and they move down until they reach a level for which they are suited. People who move down might be satisfied, or at least not deeply dissatisfied, because they have had their chance at the top and they realize that they have found their proper level. What interested me about the quotation above was that it is an extension of what is known as the "Peter principle" after the popular writer Laurence Peter (Peter and Hull 1969). Cole and Cole (1973) do not mention Peter, but instead cite as a source William J. Goode (1967), whose paper: "The protection of the inept," is delightful. I read Goode's paper to see whether, in fact, he had anticipated Peter and found that he had not. But he did have much to say about how those at the lower end of the skill spectrum are treated and I summarize his paper here.

THE PROTECTION OF THE INEPT

The Columbia University sociologist William J. Goode (1967) described a widespread social phenomenon which he called "the protection of the inept." The protection of the inept places two social needs in conflict. One need is to protect the inept; the other is to protect the group from the inept (p. 5 and p. 6). Most of the evidence presented by Goode suggests that the harm done by not protecting the inept is greater than the benefit to the group of protecting it from the inept. The issue of how the inept are dealt with is important to sociologists because the fact that the inept are generally protected in our society seems to be contrary to the frequently stated norm of achievement. Goode states his conclusion on p. 6:

> Industrial society is highly effective at production not so much because it allows the most able to assume positions of high leadership, but because it has developed two great techniques (bureaucracy and machinery) for both using the inept and limiting the range of their potential destructiveness.

Goode (1967) gives a number of examples of the ways in which the inept are protected. For example, professions refuse to divulge information about the competence of their practitioners, groups of workers low set standards which everyone can meet, and corporations do not fire their managers, they find other jobs for them. Therefore, as Goode says (p. 8):

> Few are fired for incompetence, especially if they last long enough to become members of their work group. One consequence is that, in craft or white collar jobs, higher standards are set for obtaining a job than for performance. The result is that a high level of formal education is often necessary for jobs that any average eighth-grader could learn to perform rather quickly.

One way that the protection of the inept is manifested is that members of the upper classes prevent members of the lower from developing the skills that would enable them to compete with the less talented members of the lower classes. Goode writes (p. 8):

> [E]very detailed study of a class system describes how the upper strata prevent the lower from acquiring the skills appropriate for higher level jobs. This effort alone is a good indicator that the upper strata include many who are less talented.

Goode (1967) lists a number of factors that tend to cause an increase or a decrease in the protection of the inept. Under the heading "external processes" Goode lists: (1) High demand for a service decreases the need for quality. He gives the expansion of graduate education in the 1960s as an example where a great demand was met by a low-quality supply. (2) A low supply of people willing to perform a job may lead to the acceptance of the inept. For example, top scholars do not want to be administrators, so administrative jobs may have to be filled by the relatively inept. (3) Sometimes there is a low demand for high quality, for example for pollution control [especially, I would say, when the polluters are in change of regulating pollution]. (4) Ineptitude is discouraged if it reflects badly on the boss.

Under "internal processes" are listed: (1) Having inept individuals within a group provides a floor which the merely mediocre are safely above. (2) The cost of firing the inept may be higher than the cost of their staying. (3) Constant excellence is not expected. Workers may have to take time off to tend a sick child, and some workers slow down when they get older. (4) Ineptitude may be difficult to detect. Goode gives a number of examples of this. He lists parenthood, religious behavior and administration as examples. His treatment of religion is particularly amusing (p. 13):

> The work of a clergyman is especially hard to measure, in part because there has been

> little agreement on what the performance ought to achieve. However, both his superiors and clients are more likely to measure his work by, say, an increase in church attendance rather than by the parishioners' increased rejection of sin.

The problem of measurement is a difficult one, and Goode (1967, p. 13) lists four false "answers" to the problem: "focusing on an irrelevant trait [the race or sex of an individual], seizing on an irrelevant performance [a foreman who is popular with his men but cannot get them to work], insulating members from outside competition, and barring the gateways to training." These convenient, false answers help protect the less talented or skilled.

Goode does mention a few areas in which ineptness is hard to hide: sports, law and the higher levels of basic research. He goes on to consider consequences of protecting and of not protecting the less able. The group may suffer no relative disadvantage from protecting the inept if rivals suffer from the inept as well. Goode uses the example of the military, where a man might be frustrated by the ineptness of his side, but he can be reassured that the enemy is probably just as inept. On the other hand, Goode points out that some collectives, for example, families, compete with others at the same level and protecting the inept in one family may put it at a disadvantage with respect to another family of the same social standing.

If the inept are not protected, there is a danger of the society becoming a Hobbesian jungle, Goode says. The disadvantage of not protecting the inept is that social cohesion may be destroyed. If cooperation is important in a group, rewarding some members at the expense of others may lose the cooperation of the losers and the productivity of the group may suffer.

Some structures with less protection of the inept work, for example sports teams and research departments. The less skilled members of a team benefit when the team wins, and they may be willing to subordinate their individual interests so that they can benefit from the success of the team. Researchers may work independently, but the success of one member of the group may raise the reputation of the whole group and thus benefit other members who may not be involved in the work.

Goode (1967, pp. 16-17) points out that societies vary in how they balance the protection *of* the inept and the protection of the group *from* the inept. He says that he has not found the optimal solution, but he does conclude that the modern system is as productive as it is because it can use the inept more efficiently than earlier systems.

I think that Goode's paper is interesting in itself. It is not strictly relevant to my topic of stars and fashions in science, but it is interesting that the resolution of the problem of how to deal with the inept (have them work in a bureaucracy or with machines) is what Ortega decried in *The Revolt of the Masses.* Ortega's mediocre scientists are parts of the same kind of system as Goode's modern workers.

DISCUSSION

There are four questions, or sets of questions, that should be asked about "stars" and "fashions" in science.

1. What is meant by "stars" in science?

I distinguish three meanings of the word "star" in science.

a) A "star" may be a leader in his or her field, perhaps a fashion leader. Such a leader defines what is going on in the field and may dictate what work will be done and what is acceptable. Leaders are needed for guidance and authority, but they are not necessarily needed to provide fashion advice. This is roughly the sense in which Lawrence and Locke (1997) use the phrase "star system."

b) Some academics are really stars in the same sense as athletes and movie stars. These people are media stars. Such stars are known for their personalities as well as their accomplishments. This sort of stars are needed when the press wants to tell stories and make the actors more human. This is the sense in which Shumway (1997) describes stars in literary studies. It is clear that the word "star" is appropriate in this sense, but it may not be appropriate in senses a) and c), even though it is used in those ways as well.

c) The word "star" is sometimes used to refer to someone who is merely good (or perhaps very good), who is known, or who has a good reputation. Stars, that is, people with good reputations, are needed to establish the reputation of university departments. This is the sense of the word "star" that was used by the people at the University of Minnesota-Twin Cities who were interviewed by Smetanka (1998).

2. What is the reality that the word "star" points to?

a) All fields must have leaders, for example, journal editors, research directors, and department heads. These people have formal positions. There are also informal positions occupied by people whose work is admired and who may set fashions that provide discipline for the field. The number of leadership positions, whether formal of informal, is limited even if the number of potential leaders is not. Leaders should be well-informed and have good judgment. They play a socially recognized role and have authority.

b) Some scientists are, or have been, media stars. Usually such people work in areas of great general interest, such as the nature of matter, or the origin of human life. There is much less demand for commentators on science than there is for commentators on political and social issues. It may be important to realize that while there is a need for leaders within a discipline (as in (a) above), what the media want is opinion or stories. Media stars are story-tellers and personalities. They may be experts, and their expertise may be used, but they do not have authority to decide issues of public importance.

c) Some scientists do more work and better work than others. While it is difficult to measure accurately the value of scientific work, there is a great deal of empirical work involving quantitative measures of research output and its reception.

3. What is the cause of the "star system" in science?

a) There is a need for leaders in science. There are a number of social roles for leaders in science. These people need not be stars. That is, they do not need to be outstanding, but they should be competent. Leaders are respected for the social roles they play, if not for their outstanding ability. This subject has been studied under the heading of "social stratification" by sociologists interested in the functioning of institutions.

b) The media need something to show, to write and/or talk about. The subject should be popular, or something that can be made popular. The lives of people are often of more interest than their ideas, and even if the ideas are the subject of discussion, it is convenient if the person with the ideas has had an interesting life. Journalists often have stories to tell and they want someone to quote in order to support their cases. For this they need "experts" not stars. Experts are (usually) anonymous, but they should have credentials to make them believable. Stars are people who are interesting in themselves. They are so interesting that people outside of their specialty know their names and may be interested in what they say or do.

c) Some scientists are better than their colleagues at what they do. It is clear that there are differences in accomplishment, but it is not clear what causes these differences. Is it talent, training or work opportunity? It would be interesting to know, for example, whether we could have many more skilled scientists than we have if we just trained them, or if the number of potential scientists is strictly limited by innate ability. If one were to combine the vague environmentalism that seems to pervade American social thought (environmental differences are [almost] all-important) with the observation that most scientific work is done by a very small proportion of all scientists, then one might conclude that we have about as much science as we want, and that we could have much more if we wished.

4. What are the consequences of the "star system" in science?

a) Leaders are needed to play the necessary social roles in science, but not many such leaders are needed. Worse, not many leaders are possible because of the limitation in the number of positions. Insofar as the leaders determine what work is acceptable, their interests may limit the work that gets done. They may favor their own students and friends and they may be conservative. Science is conservative by nature, but powerful old people may discourage good new ideas.

b) Media stars make science more interesting to many people, including non-scientists. They may also get the attention of their fellow scientists. But the attractions of publicity may seduce talented scientists from serious scientific work. The needs of the media may lead them to misrepresent science to make it simpler or more exciting than it is. It is important that people understand science because public support is important, but understanding is also important if the public want to understand the world. I doubt if many great scientists are lost to the profession because of their popular success. There is more danger to science if it is believed, either by practitioners or managers, that science must be popular to be worthwhile.

c) There are big differences in the productivity of scientists. One ideas, suggested by Lotka's law, as interpreted by Price (1963), is that since relatively few people publish most scientific papers, the unproductive majority might be eliminated without harming science. Cole and Cole (1972) tentatively advanced such an idea. I sometimes wonder what would happen if all of the top scientists were eliminated (I am not proposing this, but as a thought experiment, imagine that all members of the U. S. National Academy of Science were forbidden from doing research but required to teach two courses a term for the rest of their careers). The big problem is that there is strong, expensive competition among universities for the relatively small number of top scientists. Universities compete for known people. I wish that someone would ask whether what is wanted is good people or known people. The known people are good, but there can only be a limited number of people who are known. If the system of science artificially limits the number of people who have a chance to do good work, and if universities have to pay a premium for an unnecessarily rare commodity, then the institution of science is malfunctioning.

ADDENDA

There are three ways to get into the literature of stars and social stratification in science, and other literature for that matter. One way is to find a recent paper and look at the references. Another way is to find a good, old paper and look at the citations. Another way is to look up terms such as "Ortega hypothesis" or "Matthew effect" on the Internet. I think that it is likely that almost all important references can be found by working back and forth from recent papers to older papers that they cite and then forward to citing papers.

One of the most interesting statistical issues is what mechanisms produce the large variability in scientific productivity. Garfield (1980) provides a good entry to this literature. He points out that large-tailed distributions (referred to as the laws of Bradford, Lotka, Pareto and Zipf) are found in a number of areas including scientific productivity, number of times papers are cited and word frequencies. Two papers proposing mechanisms that produce such distributions are those of Simon (1957) and Price (1976).

A recent book on ecological methodology (Ford 2000) refers to some fairly recent work on citations, including work that discusses a variety of factors motivating citations. A number of useful comments are made in Ford's chapter 13 about ecological literature and scientific literature in general. A number of observations cast doubt on the norm of *universalism.*

There has been renewed interest in stars at the University of Minnesota-Twin Cities since our Governor proposed a biennial budget for the U that would not even cover inflationary increases in employee health-care costs. The problem for the U, and for other public research universities, is that they do not have the money to compete with private research universities for star faculty. A nice discussion of the recent history of competition between public and private universities was provided by Alexander (2000). A large amount of comparative data was presented at a meeting of the U of M Senate Committee on Finance and Planning (Minutes 1999). Professor Alexander was interviewed by conference call during the meeting, and among the interesting comments was this: "Legislatures must understand that the nation's greatest public institutions are becoming training grounds for private institutions." Another important point is that the system of student support used by the State and the Federal government favors high tuition at private schools (the higher the tuition, the higher the government support for students). Higher tuition at private colleges and universities, subsidized by the public, permits the private colleges to pay higher faculty salaries, and to lure star faculty from public universities. Since taxpayers are not willing to support public universities as well as they once did, tuition at the public universities must be raised if they are to try to compete with the privates.

These comments (the Addenda) were added, and the reading behind them was done, after I wrote the first part of this paper. Now I realize that some of my own ideas in this paper have already been mentioned by other authors, including Price (1976).

LITERATURE CITED

Alexander, F. K. 2000. The silent crisis: the relative fiscal capacity of public universities to compete for faculty. The Review of Higher Education 24:113-129.

Charnov, E. L. 1976a. Optimal foraging: the marginal value theorem. Theoretical Population Biology 9:129-136.

Charnov, E. L. 1976b. Optimal foraging: attack strategy of a mantid. American Naturalist 110:141-151.

Cole, J. R. and Cole, S. 1972. The Ortega hypothesis. Science 178:368-375.

Cole, J. R. and Cole, S. 1973. Social Stratification in Science. University of Chicago Press, Chicago.

Curio, E. 1976. The Ethology of Predation. Springer-Verlag, Berlin.

Editorial. 2001. 'U' stars: medical leadership points the way. Minneapolis Star Tribune, January 9, p. A10.

Emlen, J. M. 1966. The role of time and energy in food preference. American Naturalist 100:611-617.

Feller, W. 1968. An Introduction to Probability Theory and its Applications. Vol. 1, Third Edition. John Wiley, New York.

Ford, E. D. 2000. Scientific Method for Ecological Research. Cambridge University Press, Cambridge.

Garfield, E. 1980. Bradford's law and related statistical patterns. Current Contents # 10, pp. 5-12, May 12, 1980 [and pp. 476-483 in: Essays of an Information Scientist, Vol. 4, 1979-80].

Gilliam, J. F., Green, R. F. and Pearson, N. E. 1983. The fallacy of the traffic policeman: a response to Templeton and Lawlor. American Naturalist 119:875-878.

Glanz, J. 2001. The Web as dictator of scientific fashion. New York Times, June 19, 2001, pp. F1, 2.

Godfray, H. C. J. 1994. Parasitoids. Princeton University Press, Princeton, New Jersey.

Goode, W. J. 1967. The protection of the inept. American Sociological Review 32:5-19.

Green, R. F. 1980. Bayesian birds: a simple example of Oaten's stochastic model of optimal foraging. Theoretical Population Biology 18:244-256.

Green, R. F. 1986. An impecunious gambler in a kind-hearted casino. UMD Department of Mathematical Sciences, Technical Report 86-1.

Green, R. F. 1990. Putting ecology back into optimal foraging theory. Comments on Theoretical Biology 1:387-410.

Gross, M. R. 1994. The evolution of behavioral ecology. Trends in Ecology and Evolution 9:358-360.

Heilbron, J. L. 1977. The upward path (a review of Harriet Zuckerman's book, Scoientific Elite: Nobel Laureates in the United States.) Science 196: 754-755.

Hinde, R. A. 1982. Ethology: Its Nature and Relations with other Sciences. Oxford University Press, Oxford.

Krebs, J. R. and Davies, N. B. 1978, 1984, 1991, 1997. Behavioural Ecology. Blackwell Scientific Publications, Oxford.

Lawrence, P. A. 1999. The fashions and foibles of science (a review of the book, Egg and Ego: An Almost True Story of Life in the Biology Lab). Nature 397:487-488.

Lawrence, P. A. and Locke, M. 1997. A man for our season. Nature 386:757-758.

Lima, S. L. 1984. Downy woodpecker foraging behavior: efficient sampling in simple stochastic environments. Ecology 65:166-174.

Lima, S. L. and Dill, L. M. 1990. Behavioral decisions made under the risk of predation: a review and prospectus. Canadian Journal of Zoology 68:619-640.

Lloyd, M. 1967. Mean crowding. Journal of Animal Ecology 36:1-30.

Locke, M. 1997. The decline of universities with the rise of edubis. Pp. 8-16 in: R. W. Nelson, (ed.) Inside Canadian Universities. Cedar Creek Publishers, Kingston, Ontario.

Lotka, A. J. 1926. The frequency distribution of scientific productivity. Journal of the Washington Academy of Science 16:217-223.

MacArthur, R. H. 1972. Geographical Ecology. Harper and Row, New York.

MacArthur, R. H. and Pianka, E. R. 1966. On optimal use of a patchy environment. American Naturalist 100:603-609.

Mangel, M. and Clark, C. W. 1986. Toward a unified foraging theory. Ecology 67:1127-1138.

Mangel, M. and Clark, C. W. 1988. Dynamic Modeling in Behavioral Ecology. Princeton University Press, Princeton, New Jersey.

McNamara, J. M. and Houston, A. I. 1986. The common currency for foraging decisions. American Naturalist 127:358-378.

Minutes of a meeting of the [University of Minnesota] Senate Committee on Finance and Planning, Tuesday, November 2, 1999.

Nelson, C. 1997. Superstars. Academe, January-February pp. 38-43, 54.

Oaten, A. 1977. Optimal foraging in patches: a case for stochasticity. Theoretical Population Biology 12:263-285.

Ortega y Gasset, J. 1932. The Revolt of the Masses. W. W. Norton, New York.

Peter, L. J. and Hull, R. 1969. The Peter Principle: Why Things Always Go Wrong. William Morrow & Company, New York.

Price, D. J. D. 1976. A general theory of bibliometrics and other cumulative advantage processes. Journal of the American Society of Information Sciences 27:292-306.

Price, D. J. D. 1963. Little Science, Big Science. Columbia University Press, New York.

Pyke, G. H. 1984. Optimal foraging theory: a critical review. Annual Review of Ecology and Systematics 15:523-575.

Pyke, G. H., Pulliam, H. R. and Charnov, E. L. 1977. Optimal foraging: a selective review of theory and tests. Quarterly Review of Biology 52:137-154.

Schoener, T. W. 1971. Theory of feeding strategies. Annual Review of Ecology and Systematics 2:369-404.

Schoener, T. W. 1987. A brief history of optimal foraging theory. In: Foraging Behavior. (Kamil, A. C., Krebs, J. R. and Pulliam, H. R., eds.) Plenum Press, New York, pp. 5-67.

Scott, J. 1997. Scholars fear 'star' system may undercut their mission. New York Times December 20, 1997 pp. A1, B9.

Shumway, D. R. 1997. The star system in literary studies. Publications of the Modern Language Association 112:85-100.

Sibly, R. M. and Smith, R. H., eds. 1985. Behavioural Ecology. Blackwell Scientific Publications, Oxford.

Simon, H. A. 1957. Models of Man: Social and Rational. Chapter 9: On a class of skew distribution functions. John Wiley and Sons, New York.

Smetanka, M. J. 1998. 'U' science makes some feel short-shrifted. Minneapolis Star Tribune, Febuary 21, pp. A1, A9.

Stephens, D. W. and Krebs, J. R. 1986. Foraging Theory. Princeton University Press, Princeton, New Jersey.

Templeton, A. R. and Lawlor, L. R. 1981. The fallacy of averages in ecological optimization theory. American Naturalist 117:390-393.

Turelli, M, Gillespie, J. H. and Schoener, T. W. 1982. The fallacy of the fallacy of the averages in ecological optimization theory. American Naturalist 119:879-884.

Ziman, J. M. 1984. An Introduction to Science Studies. Cambridge University Press, Cambridge.

Zuckerman, H. 1977. Scientific Elite: Novel Laureates in the United States. Free Press, New York.