"Interview of J. Scott Armstrong" for the International Journal of Forecasting, By Fred Collopy, Case Western Reserve University

The Oracle would like to thank the *International Journal of Forecasting* for granting permission to publish this interview in the Oracle. An "Interview of J. Scott Armstrong" by Fred Collopy is to be published in a forthcoming issue of the *International Journal of Forecasting*. In the interview, ForPrin.com founder and Director Scott Armstrong describes his life in forecasting from before the founding of the International Institute of Forecasters and up to the present day.

I have known Scott Armstrong for a quarter of a century. He has been my teacher, collaborator, and friend. He is an iconoclastic thinker, a passionate and persistent scholar, and a warm human being. He is one of only two members of the International Institute of Forecasters (IIF) who has attended every single International Symposium on Forecasting. His contributions to the Institute, the two journals, several popular websites and the fields of forecasting and marketing are too numerous to keep track of (although he does a pretty good job of it on his websites).

While Scott has been honored as a Fellow by the IIF and is well known in the forecasting community, we are overdue in hearing the story of his contribution to and thoughts about our collective enterprise. With that in mind, I was pleased when Rob Hyndman asked me to interview Scott for the International Journal on Forecasting.

We conducted this interview by email.

Scott, before we get into your background and reflections on forecasting and your research, I would be interested to hear any thoughts you have about this email interview approach.

I like it. It will give me a chance to reflect on your questions and my life as a forecaster.

Ok, so to get started, how did you become interested in forecasting?

My work experience played a role. One of my first summer jobs was as an asbestos worker. My job title was "Improver". I liked that title and decided to keep that role. If my work is not going to improve things, why bother?

After graduating from Lehigh as an Industrial Engineer in 1960, I went to work for Eastman Kodak in their Print and Processing Division. Each day, someone would get on the phone to tell a few workers whether or not to stay home the next day, as there might not be enough work. I thought there must be a way to predict how many workers would be needed further in advance. This was about the time that Robert G. Brown was publishing his work on exponential smoothing. It was a thrilling experience applying his methods. I collected the data and made adjustments, wrote a program, had it punched on IBM cards, and carried the heavy boxes of cards to a room with a giant computer. It then labored for hours to produce results by the next morning—something that would be nearly instantaneous on a laptop today.

The exponential smoothing model was more accurate than the accountant, Jim, who plotted the data, checked the weather forecast, and then came up with his unaided judgmental predictions. Kodak did not want to adopt my system while Jim was around; however, they adopted the program a couple of years later when he retired. I had started on my journey as an improver.

As an aside, because computers were so slow in those days, I learned that one needed to assess prior research carefully before analyzing the data. This is still useful. The current reliance on statistical programs is unfortunate. Forecasters and practitioners believe that statistical analyses of non-experimental data will enable them to discover relationships—a belief which has led to my crusade against such methods as stepwise regression, factor analysis, and data mining, This started with my "Tom Swift studies" on regression and factor analysis. Tom Swift would obtain statistically significant results by using standard statistical procedures on, as it turned out, random numbers (Armstrong 2012).

After my experience at Kodak, most of the problems that I encountered in my career involved forecasting, if only for testing validity. For example, one year after leaving Kodak, when working for a small company which had recently changed its name from Haloid to Xerox, I used exponential smoothing to develop a forecasting system for inventory control. In that case, implementation was easy because there was no prior forecasting system at Xerox.

I continued my interest in forecasting while earning my MS (now MBA) at Carnegie Tech (now Carnegie Mellon). It was there that I worked with John Farley, a marketing professor, testing whether Markov chains, which were popular at the time, were useful for forecasting. They were not. This reinforced another of my basic ideas: that the primary task of management scientists is to determine which methods are most effective in a given situation.

As I was leaving Carnegie Mellon in 1965 to go to the Sloan School at MIT to do a PhD, John Farley gave me two pieces of advice: (1) try not to take any courses, and (2) do not work in areas where there are a lot of other researchers. This was excellent advice.

However, I immediately violated John's first piece of advice. At the Sloan School, PhD candidates could create their own program. Because I viewed myself as an economist who was interested in practical issues, I decided to spend my first semester taking courses in their highly rated economics department. It was a dreadful semester; the courses violated my understanding of logic, my need for evidence, and my value of freedom. The faculty thought their job was to tell others how to live. To support this, they relied on the opinions of Keynes who, like Marx, had little regard for scientific evidence. They spoke in math, rather than in plain English. After that semester, I stopped taking courses and took full advantage of the opportunities offered by MIT. The design of their PhD program was ideal for learning, as well as for doing research. We could develop our own objectives, then prepare a course of study and a thesis, and find faculty who would agree that the plan was worth supporting. It has been sad to see universities gradually departing from such a learner-responsible model over the years (Armstrong 2011).

In my search for important problems for my thesis at MIT, I considered three possible topics, all of which involved forecasting, namely happiness, crime, and the one which I chose: "long-range forecasting for a consumer durable in an international market." That, plus ten years of additional work after the thesis was accepted, led to my first book, Long-Range Forecasting, in 1978.

"An Interview with J. Scott Armstrong," contd.

I was especially interested in forecasting because it allowed me to work on important problems, and to be an improver, because people often use inappropriate forecasting methods. They are inappropriate because experts wrongly believe that they can rely on their own judgments to make accurate forecasts. I addressed this in my seer-sucker theory paper (Armstrong, 1980), and it has since been supported by Phil Tetlock's (2005) landmark study. Unaided expert forecasting is useless for complex situations involving uncertainty. This finding is hard for people to accept; think about economists, climate scientists, CEOs, and political leaders.

Legal cases provide important problems for forecasters. I was fortunate when in the early 1970s Ford Motor Company asked me to be an expert witness and forecast lost profits for a defunct Philadelphia dealership. Ford had lost two previous cases and did not want to lose this one. The dealers had enlisted a high-priced consulting company to prepare forecasts of enormous losses. I was excited to learn this because of my belief that few management consultants know much about forecasting. It was, as they say, like shooting fish in a barrel.

To date, I have been involved in 13 legal cases, and my clients have won all of them. The reasons, I believe, are that the opponents know little about forecasting, and that I only accept cases where, in my judgment, my clients are in the right. I found lawyers to be a delight to work with. They tell me what they want then give me complete control over spending, provide no bureaucratic obstacles, and agree with my contract, which is that they keep me fully informed and do not try to influence my testimony. I have encountered only one law firm that violated this contract by withholding key information. I am currently involved with a paper which arises out of my expert testimony on the effects of government-mandated advertising disclaimers. Short story: they infringe on free speech with no demonstrable benefits (Green and Armstrong 2012).

The not-for-profit sector, especially in the government, is another source of important forecasting problems. (In the for-profit sector, prices serve as forecasts and profits as an objective measure of success or failure.) I have been involved in a number of projects in the government sector, and have found that the forecasting methods used range from unaided judgment to mindless statistical procedures. My first such project occurred during the Cold War: I was asked to forecast the health of political leaders around the world to see how that might affect their relationships with the U.S. (I was not told who provided the funding, but you can guess). I have since worked on forecasting related to terrorism, the development and use of methods of mass destruction, and, in 2006, a 75-year-ahead forecast of U.S. Medicare expenditures for the U.S. Congress. More recently, I have been fortunate to have been able to work with Willie Soon, of the Harvard-Smithsonian Center for Astrophysics, and with Kesten Green on climate change. This led to the opportunity to testify before a U.S. Senate committee in 2008 regarding forecasts in relation to the endangerment of polar bears, and before a U.S. House committee in 2011 on the validity of global warming forecasts (they are not valid). These were times when I thought, "My parents would have been proud of me!"

That 1978 book, Long-Range Forecasting, was where I first encountered your work. There was so much about that book to distinguish it—lots of empirical results, an extensive annotated bibliography, even the occasional cartoon. But perhaps the most irreverent or unconventional feature was the Don't List, a list of articles on forecasting with promising titles that you considered a waste of the reader's time. How did your editor, readers, and other researchers respond to your directness?

Let's start with your first point about "empirical results". My objective was to examine all useful research findings and to translate them into advice for researchers and practitioners. I had naively assumed that this evidence-based approach was the basis of the scientific method. But while Long-Range Forecasting (LRF) had highly favorable reviews in general, there were rumblings from management scientists who felt that LRF contained attacks on their work. Indeed, the evidence often conflicted with common practice and popular theories.

Much of the mainstream research in forecasting did not yield any useful findings. Some of the more prominent such papers were included in the Don't List. It contained 160 books and papers that one did not need to read after reading LRF. It said, "If you do not read each and every one of them, you will save at least 500 hours". Some of these papers were discussed sufficiently in LRF, and some had been superseded by more recent papers. For example, one of my own papers was on this list. Many papers were there because the titles looked so promising, yet they had little to offer the target audience. I thought that the list would save time for people because journal abstracts typically fail to describe the findings or how they were obtained. Researchers did not seem to be offended by the Don't List. In fact, two papers by Spyros Makridakis, whom I did not know at the time, were on that list. After reading the book, he contacted me and invited me to visit him at INSEAD. I did, and we have had a long and fruitful relationship.

That relationship with Spyros, along with Robert Fildes and Bob Carbone, was responsible for launching the International Symposium on Forecasting in 1980, the two journals (the Journal of Forecasting and the International Journal of Forecasting), and the Institute. Would you share that story with us?

Publishing research on forecasting was particularly difficult because, as a discipline that spanned conventional subject areas, it had no obvious "home" in university departments. Moreover, mathematics and theorizing ruled the day, and research which was based on the empirical testing of alternative approaches in realistic situations was difficult to get past reviewers. Fortunately for me, and for the discipline, Spyros Makridakis and Robert Fildes shared an interest in "evidence-based forecasting" and in studying important issues. Along with Robert Carbone, we started the Journal of Forecasting (JoF) in 1981. The "Aims, Scope and Standards" section stated: "Of particular interest are papers that compare different methods in actual situations".

I offered to write the standards for the journal. I had developed an interest in this area, because my most useful evidence-based findings were being rejected by journals, often with quite nasty reviews—a problem which has worsened over the years. This led to my interest in peer review. This research, summarized in the first issue of the Journal of Forecasting (Armstrong, 1982), led to what were, at the time, innovative peer review procedures for our journal.

Focusing on the testing of alternative methods—as demonstrated by the ground-breaking work related to the forecasting competitions conducted by Spyros and Michele Hibon starting in the 1970s—the JoF was an immediate success. By 1983 it had the sec-

ond highest journal impact factor of all management journals. This was no accident: it was a result of our use of peer-review procedures that were supported by research. Few journals used such procedures then, and most editors continue to rely on their opinions and tradition for their reviewing practices.

In an interview for the IJF, Spyros said that the four of you who started the Institute have contributed a lot to the field "each in his own way." Would you characterize some of those contributions, your own and those of the others?

Spyros was the one who proposed starting a journal, and he had a good plan to make it a success. He obtained a contract with the publishers John Wiley & Sons. He put a team together and asked me to work on the editorial policies. Spyros believed that the focus should be on publishing important research. We believed that reviewers should be used primarily to improve papers, not to decide what should be published.

We were a creative and volatile group of four. We were fortunate to have Robert Fildes to guide the process and to hold things together during those early years. Robert continues to perform this role in his own inimitable way.

Given that there were only four of us, decisions were made quickly, and this helped us to put our policy of publishing useful papers into action. For example, the reviewers rejected a paper that Spyros thought was useful, so we published it (Lawrence, Edmundson & O'Connor, 1985). Spyros was right: it proved to be a well-cited and influential paper. The flip side of the policy was that we received many papers that were done carefully by competent authors, but were of no obvious value, so we did not send them out for review. Our requirement for usefulness trumped the concept of "fairness" in our decisions.

Our strategy for publishing useful research was to invite good researchers to publish papers on topics that were relevant to our aims. This was Spyros's initiative, and I was a bit skeptical at first. Most of those who we invited accepted, and we published the papers when the authors said they that were ready. As we found in a follow-up paper (Armstrong & Pagell, 2003), this strategy was 20 times more likely to result in important papers than traditional channels.

Bob Carbone had the idea for the annual symposium and ran the first one, a very successful International Symposium on Forecasting (ISF) in Quebec City in 1981, with 450 participants. I assisted Bob and learned a great deal from him about how to run a conference. That led to a most satisfying experience when I was General Chair and Program Chair of the Philadelphia conference in 1983, which drew 1,100 participants.

An international band of researchers who were dedicated to the testing of alternative approaches joined us at the ISFs. My wife, Kay, and I began to build vacations around the annual ISFs, which were held in various international venues.

We welcomed new ideas at the ISFs. To help ensure this, we did not use peer review to decide who could present their research. Instead, we regarded it as our duty to allow everyone who wanted to do so to present their research. We also encouraged participation by practitioners. For example, in Philadelphia about half of the attendees were practitioners. This was all part of our plan to try to bridge the gap between academic researchers and practitioners.

On the strength of our early success, we asked our publisher, John

Wiley, to extend the contract on a basis which was more favorable for us. We made a poor forecast about the outcome of the negotiations, which led to a split with Wiley. As a result of the failed negotiations, we subsequently founded the International Journal of Forecasting, published by Elsevier. Fortunately this wasn't the end of amicable relations: our Wiley editor, Jamie Cameron, was a great help in the success of our early efforts, and Jamie and I remain good friends to this day.

Wisely, Robert and Spyros reasoned that the long-term success of our efforts required that we found an institute, the International Institute of Forecasters, and invite others to participate. This produced many benefits. For example, the ISFs would never have survived without the many people from all over the world who have volunteered to run them. The 31st ISF, which was held in Prague in June 2011, is a testament to the strength of the Institute.

The Institute was also instrumental in the founding of Foresight, thanks to Len Tashman. Foresight has been crucial in our efforts to bridge the gap between academics and practitioners, and to pick up on some of the original aims of the Journal of Forecasting, which stated that it should include: "teaching materials, notices of general interest, annotated bibliographies to relevant research published elsewhere, a practitioners' forum, letters to the editor, and reviews of research."

Perhaps inevitably, by expanding the IIF Board and opening up editorial positions, the Institute became more like other academic institutes and journals. For example, as with other journals, the concept of "fairness" has crept into the reviewing process for the JoF and IJF. My unwillingness to compromise on our original aims—notably the JoF aim "Theoretical contributions aimed at a narrow audience, studies merely fitting models to past data, or work that does not deal with important subjects should not be submitted to the Journal"—conflicted with the desire of new board members to make our journals more like those in other disciplines and to make the journal consistent with the reward systems imposed by universities. In any event, around 1990, the Board apparently decided that they no longer needed my advice as Editor or Director.

How did you react to that?

It was disappointing. However, it turned out well for me. Most importantly, it freed up a lot of my time. This was put to good use when I met other researchers who had similar views about research.

The first of those was you, Fred. We began to work together around 1986 when you were working on your PhD at Wharton. Our work on rule-based forecasting, forecast evaluation, and the relationship between judgmental and quantitative methods was an exciting time. Your focus on condition-action statements as a way of transmitting knowledge was one of the most useful things that I learned from you. It dictated the design of my books Principles of Forecasting (2001) and Persuasive Advertising (2010).

In 2000, Kesten Green, then working in New Zealand (now in Australia) and I began to collaborate on forecasting for conflicts (e.g., terrorism and wars). We developed and tested the "simulated interaction" and "structured analogies" methods. They can help in finding effective ways of resolving conflicts. We found, for example, that simulated interaction would have predicted correctly the unfortunate outcome of the strategy we used in the above-mentioned conflict with John Wiley.

Starting in 2007, Kesten and I began to analyze the forecasting methods used by global warming alarmists. In effect, climate change is a forecasting problem. We have concluded that there are no scientific forecasts that long-term warming will occur, or that warming would have net negative effects if it did occur, or that any government policies would help if there were in fact dangerous manmade global warming. In a current paper, we conclude that the global warming alarm is an anti-scientific political movement, much like the global cooling movement of the early 1970s, and at least 25 other movements that have suggested that mankind is causing harm to the Earth (Green & Armstrong, 2011).

In early 2004, Alfred Cuzán, a political science professor at the University of West Florida, contacted me with ideas about how to introduce policy variables into political forecasting. That led us, along with Randy Jones, a political science professor at the University of Central Oklahoma, to launch PollyVote.com and the political forecasting Special Interest Group at forprin.com. We use these sites to demonstrate the value of the forecasting principles, such as combining forecasts. The PollyVote, our combined forecast, is now so accurate that it might be more accurate than the official vote counts in the U.S.

In 2007, I met Andreas Graefe, and we found that we had many common interests. He spent a year working with me at Wharton while he finished his PhD at Karlsruhe Institute of Technology in Germany, and the following year on a post-doc fellowship. We continue to work together, primarily on the use of the "index method" for forecasting in situations with many important variables and a lot of knowledge about the effects of the variables—such as, for example, in political elections and other selection problems. The research can be used to aid decision making, such as which candidate to nominate to stand for elected office and which issues to emphasize. Andreas is now the creative force behind PollyVote.com.

In addition, I have still remained an Associate Editor of the IJF, and continue to offer advice in that role. Of course, I also make regular contributions to the Symposia.

You spoke earlier about the IIF's desire to reduce the gap between academics and practitioners. What are your thoughts about that gap now? Does the field of forecasting represent a good model of how academic research can assist practitioners?

Compared to other fields, we have been successful. However, while the International Journal of Forecasting does better than most journals with respect to publishing empirical tests of reasonable alternative approaches to important problems (useful findings), these form only a small percentage of the published papers. Armstrong and Pagell (2003) estimated that, across all journals, such forecasting papers have appeared at the rate of about one per month. Interestingly, over half of these papers appear in the IJF (about 5 papers per year) and JoF (2 papers per year).

I am disappointed by the reaction of scientists to evidence-based forecasting. Researchers often suppress new and useful findings under the guise of protecting the public; they call the process peer review. Imagine what would happen if companies tried to protect their customers against information about useful new products.

Journal peer review is especially distressing for improvers. Nobel Prize winners have typically said that they had difficulty in publishing their best work. Economist Julian Simon wrote about his poor treatment by journal reviewers; so he made his biggest impact through

his books. When I analyzed what I thought to be my 20 best papers, all of them had at least one negative review, and few reviews were constructive. The typical review time for these papers was five years. Fortunately, a guardian angel came to my rescue—someone I hardly knew at the time. His name was Gary Lillian, a professor at Penn State; and he surmised that my work was being treated poorly by reviewers. As Editor of Interfaces, he made me a contributing editor so that I could publish my work and work by other improvers. We called this the "Ombudsman column." This made my life easier, and I have continued in the role for many decades, thanks to the Editors who followed Gary. Some of these papers have received a considerable amount of attention, such as "Reaping benefits from management research" (Armstrong & Pagell, 2003), which found that, in forecasting journals, special treatment by editors (such as invited papers) produced papers that had 20 times as much impact as those accepted through the traditional review system—and special treatment papers are much less costly for both authors and editors.

Researchers often ignore new findings. For example, Fildes and Makridakis (1995) found that statisticians paid little attention to the findings of the M-Competition as to which statistical procedures actually work.

Perhaps the most vivid example of scientists' resistance to evidence-based findings relates to our work on climate change. People have refused to provide us with the data they used in their papers (data that had apparently been collected at the public expense), failed to provide sufficient information about their procedures, and refused to respond to questions that we asked about their research. In addition, they have removed materials from my Wikipedia entry, and used vicious language and ad hominem arguments to disparage our findings. For example, the Nobel Prize winner, Paul Krugman, implied in his New York Times column that it was absurd that I had been asked to testify before a congressional committee, as I was merely a marketing professor.

My biggest disappointment is with software providers. Software is the ideal way of introducing innovations because people are unlikely to opt out of using the most effective procedures; yet software developers are slow to adopt new findings. Most of them fail to take advantage of Ev Gardner's landmark work on damping (Gardner & McKenzie), or of the findings on combining forecasts, despite the enormous gains that have been shown from such techniques (summarized most recently by Graefe, Armstrong, Jones, & Cuzán, 2012). Nor have they included many of the findings from rule-based forecasting, such as the contrary-series rule, which states that one should not extrapolate a historical trend if there is a causal expectation that the trend should be in the opposite direction.

ForecastPro and SAS have been among the leaders in keeping up with the research, but even they could do much more. One of their key people told me that they would respond if their clients asked. Practitioners should ask software providers to include the latest evidence-based procedures. The Forecasting Selection Tree and the Forecast Audit software are tools that can help practitioners to do this. I suspect that they will encounter many providers who are unaware of many of the useful forecasting procedures available.

Academics can also help by testing the software packages. In one case, a software provider became upset at the results of a published study which found that the company's forecasting methods were

inferior to other methods it was compared with. The company filed a suit against this academic's university—but backed off when confronted by the University's lawyers.

Customers are partly to blame, as they do not insist that software be up-to-date. They could use the forecasting audit software and the ForPrin.com site to ask reasonable questions such as "Do you use Gardner's damped trend procedure?" "Do you use the Miller-Williams damped seasonal adjustment procedures?" "Do you use the contrary series rule?" "Do you combine forecasts, and if so, how?" "Do you use successive updating on hold-out samples to test the accuracy of the methods?" Few providers will be able to answer these questions in the affirmative, and many will not even understand the questions. I advise customers to avoid software that does not provide full disclosure.

I do not think formal courses hold much promise for teaching people how to forecast properly. For a start, most courses which I have seen do not use an evidence-based approach. Furthermore, few students seem motivated to learn. Also, I do not expect business schools to change. The typical approach in our business schools is to penalize those who depart from the standard approach. I have only met one administrator of an educational program who was interested in trying new approaches. That was Lars Wiberg, Director of Executive Education at the Stockholm School of Economics, and we spent well over a year developing self-learning exercises that required the students to take responsibility for their own learning. I have summarized past research on learning, and it seems clear that learning occurs more rapidly when students take individual responsibility (Armstrong, 2011). This is a basic assumption in the design of the forprin.com site. We want to provide knowledge and exercises to enable practitioners to teach themselves to use evidence-based methods and principles.

Who have been your major intellectual influences?

My 9th grade chemistry teacher took me aside one day and said that I should become a scientist. I was surprised at the suggestion. After all, I was not even fond of chemistry. Perhaps it was my interest in seeking evidence for conclusions.

I can remember only one course from my undergraduate education at Lehigh that had a strong influence on me. It was the first class that I walked into as a freshman and it dealt with logic. I wound up tutoring others in the class, including our "Little All-American" football quarterback. No way could the school have risked him failing a course.

Perhaps the most influential paper that I read on the scientific method was that of Chamberlin (1965). He compared sciences that were successful with those that were not, and concluded that the successful ones relied on the experimental testing of multiple reasonable hypotheses. Kealey (1996) offers historical support for this viewpoint, and Sinclair Lewis' 1925 novel, Arrowsmith, also reinforced this view, and had a big impact on me.

Evidence, logic, and the experimental testing of multiple reasonable hypotheses are not the complete picture for me. My research design is also influenced by the value I attach to freedom. I believe that consenting adults should be free to arrange their lives and make agreements as they see fit, as long as they do not cause substantive harm to others. I think that freedom is a more important goal than material wealth, though they usually go together. Various discussions with Ed Prescott, my roommate at Carnegie-Mellon in 1963 and

winner of the 2004 Nobel Prize in economics, have strengthened my libertarian beliefs.

My beliefs have influenced my approach to public policy issues, primarily by broadening the question. Instead of assuming that there is a problem and that the government can solve it, I ask, "Will there be a problem, and if so, how might it best be solved?" I also regard freedom as a benefit, and its loss as a cost. Governments should not act unless there are published and replicated independent forecasts from fully described scientific forecasting methods which show convincingly that a harm will occur and that the proposed actions will result in net benefits which are substantially greater than those from other feasible actions, including no government action. In our study of the global warming alarm, we have found that no such forecasts exist. Instead, governments, interest groups, and big business have used scare tactics in order to extract, respectively, control, preferred policies, and profits. My efforts in this area are, I think, more important than those in any other project that I have been engaged in.

Julian Simon was my role model as a researcher. I met him in 1980 when I was being interviewed for a chair in marketing at the University of Illinois at Urbana-Champaign. We kept in touch over the years, discussing research and planning a book and a paper, but we never did get around to them. His approach was to pick important problems such as population control and immigration, use evidence to test the alternative approaches, be objective, state the conclusions clearly, provide full disclosure, and be persistent. I viewed him as an improver who had a great influence, such as through his introduction of auctions for airline seats.

Julian had a five-year head start on me in life, so he was able to share his wisdom when I encountered critics who were upset by my research findings. He died in 1998, but I still hear his voice and feel that he is helping me. His 1980 bet with Paul Erlich on whether the world would run out of resources served as a model for my climate bet with Al Gore (see theclimatebet.com). I dedicated my book Principles of Forecasting (2001) to Julian. Julian was also one of the first to identify global warming as an anti-scientific political movement. Kesten Green and I have found support for his viewpoint in our current research (Green & Armstrong). It seems to me that Julian and I have lived parallel lives. This impression was reinforced recently when I read his autobiography. Like me, he met many people who were upset by his efforts to improve things.

Which developments in forecasting will have the biggest impact over time?

My career goal has been to find the areas that are likely to have the greatest impact and to get involved with them. Thus, my research projects reflect my answer to this question.

Most important decisions are made judgmentally. They involve such things as how to react in conflict situations, who to marry, who to hire as a CEO, whether the US military should intervene in other countries, and whether to have a medical operation. In recent years, there have been substantial gains in techniques that can contribute to judgmental forecasts. Kesten Green and I have been involved in this area. We developed and tested simulated interaction, a version of role-playing. We have also introduced and tested structured analogies.

Another key area in forecasting is the integration of judgmental

and quantitative methods. My hope is that rule-based forecasting (RBF) will come to be used more extensively for this purpose, although it currently lacks a good software program for its implementation. It summarizes what has been learned about forecasting and relies on domain knowledge from experts. It can also be used to test whether new techniques make contributions to what is already known. The rules were fully disclosed in our paper (Collopy & Armstrong, 1992).

Still another important area is forecasting in situations with many important variables and much prior knowledge. Andreas Graefe and I have been building on ideas from Benjamin Franklin, early 20th century practitioners, and on research from the 1960s on "unit weights" in order to develop what we call "index methods." Index methods can use all empirical knowledge on the issue to decide which variables are important and what the directional effects are. They are designed primarily for selection problems. We have used them, for example, to determine which candidates should run for political office and what issues they should focus on.

The implementation of new methods is critical. I expect that eventually a software provider will take the lead to build on the research that has been developed in the past few decades.

When the defaults provided by software are based upon the best available research findings, they will get used in practice.

The use of a forecasting audit can help organizations to identify new forecasting procedures. This was the reason why I, with the help of about 40 experts, developed the Forecasting Audit program. I am hopeful that organizations, consulting firms, law courts, and the various parties involved in the assessment of public projects will come to use the Forecasting Audit software on the ForPrin.com site. For example, Kesten Green and I used the Forecasting Audit to reach our conclusion that the global warming movement violates 72 of the 89 principles relevant to this situation (Green & Armstrong, 2007). We have also established publicpolicyforecasting.com to encourage others to publish forecasting evaluations of public policy alternatives.

Of course, this answer represents my unaided judgmental forecasts, and, as I noted in my seer-sucker theory, such forecasts are notoriously inaccurate. However, in my "Quarter-century review" (Armstrong, 1986), I identified only one area as having "excellent" prospects for useful research: "expert opinion." The gains in this area have exceeded my expectations. I identified two other areas in 1986 as having "good" prospects: "decomposition" and "uncertainty," and indeed, useful research has been done in these areas in the most recent quarter-century.

Are you optimistic about the future of the field?

I am optimistic. My career in forecasting has been rewarding, especially because it has brought me into contact with many researchers who also believe that progress depends on testing alternative forecasting methods. There are too many such researchers for me to list here, but you evidence-based forecasters know who you are.

The International Journal of Forecasting continues to be successful, thanks to having many capable and dedicated editors who have followed in the footsteps of Spyros Makridakis and Robert Fildes. In recent years, Rob Hyndman has introduced innovations to improve the reviewing procedures and ensure that proper disclosure is provided, and so on. He has also put more emphasis on special issues. His efforts have pushed the IJF to its highest ever Citation Impact Factor.

I am especially pleased by Rob's efforts to strengthen replication, as it is the backbone of science. Replications often conflict with the original studies in important ways (Evanschitzky & Armstrong, 2010). They are also the best way of dealing with the apparently increasing issue of cheating on the part of scientists. Incidentally, my advice is that one should never accuse a scientist of cheating; replication is more convincing, and whistle blowers often lose their jobs (Armstrong, 1983).

Interestingly, research on peer review continues to reveal ways to improve the system (Armstrong, 1997). In view of this, I would urge even more invited papers, as well as a reviewing system that focuses on improving papers. This was the strength of the JoF when we started it in 1980. I also hope that the IJF will start to publish reviews along with the papers (via the website)—and allow reviews by all who care to post civil, signed submissions. I think that the IJF should take the lead and publish all papers that are submitted that involve the empirical testing of alternative approaches to important forecasting problems—along with attributed reviews. And more meta-analyses should be invited.

If all submitted papers were published, the number of papers would drop substantially. Why? Because it would become clear that it is senseless to reward the faculty for merely publishing papers. Only those who have useful contributions would bother to publish. My estimate is that fewer than 5% of all papers published in leading journals are useful, and many colleagues agree. In addition, this procedure would be much cheaper for both authors and editors.

While I reject the idea that reviewers should decide what I am allowed to read in journals—or publish in them—I am a strong advocate of peer review as a way of improving papers. Frey (2003), approaching the topic from the viewpoint of basic economics, reached the same conclusion. To ensure that a paper is free of substantive mistakes, many reviewers, not just two or three, are needed. An experiment by Schroter et al. (2008) found that individual reviewers caught only about 30% of intentionally introduced errors. This implies that one would need about ten reviewers in order to be confident that all major errors have been found.

For me, one of the most important peer review procedures is to contact all researchers who I have cited in substantive ways to make sure that I have properly summarized their findings and that I have not overlooked key studies. Mistakes in summarizing findings from cited studies are common in the scientific literature (Wright & Armstrong, 2008).

Peer review is especially useful for improvers, as their findings are challenged. I am thankful to the many reviewers who have helped me with my research. For example, 123 reviewers helped with my Principles of Forecasting book. In particular, my wife, Kay has helped on all of my major papers and books. I make many mistakes in my research and writing. That is why my book Persuasive Advertising went through 272 versions before it was submitted. Fortunately, my reviewers catch the errors. I am only aware of one serious error which was not caught: I had concluded that scenarios could lead to improved forecasts, while research has actually shown that they produce misleading forecasts.

It is not enough to inform practitioners about new methods that have been shown to improve forecasting; they must also be easy to

apply. Unfortunately, as mentioned above, software providers seem unaware of or reluctant to use this useful knowledge. To overcome this, I would like to see the IIF sponsor the development of software for implementing fully disclosed evidence-based methods. This would allow organizations to use the software at little cost.

The primary reason for my optimism is that technology will eventually make the peer review system obsolete as a way of controlling what is published. The Internet allows us to communicate new findings, provide open and continuous peer review, and make the cumulative knowledge easily available to practitioners in an understandable way. Our contribution to this effort, supported initially by the Wharton School, and currently by the International Institute of Forecasters, began in 1998 with the founding of forecastingprinciples.com (forprin.com). It took ten years to reach a million visitors in total. Thanks to the leadership of Kesten Green, the site has grown enormously, and is expected to attract almost a million visitors in 2012 alone. That exceeds my expectations.

Technology, especially the Internet, is also useful in that it makes it easier to provide full disclosure of forecasting projects. This is critical for government projects. For example, I was responsible for forecasting for a "mini-car mass transit" project in 1970. This was an expensive project funded by the U. S. Department of Transportation, with federal money going to the University of Pennsylvania, General Motors, and a start-up firm. My research showed that, contrary to the goals of the project, the system would increase social costs because more people would be enticed to switch from mass transit to the mini-car system than vice versa. Strangely, my reports kept getting lost. When I was persistent that we incorporate the findings into the recommendations, the project personnel (a large group) held a meeting to have me removed from the project. A number of researchers supported me so that that did not happen. It occurs to me now that had the project been transparent, by being on the Internet, it would have ended much sooner. We might also have posted communications from interest groups that tried to sway our findings. Interestingly, our research showed that there would be a demand for the service, and in 2002, five young community activists, with little financial backing, started Philly Car Share. It has been successful and has been emulated in cites around the U.S.

I have found my career as an improver immensely rewarding. I love to get up every morning, including on vacations, to work on my latest project. However, it is only fair to warn you that I, like Julian Simon and others in the academic world, have to pay a price for being an improver. It is amazing how upset people become when your findings do not agree with their opinions. You can get an idea of the life of an improver by reading Julian Simon's autobiography, A Life Against the Grain.

I keep a score sheet for my improvements in forecasting in terms of the number of discoveries in forecasting that I have been involved with: 25 so far, usually with other researchers. I have been listing these discoveries on my CV on jscottarmstrong.com for as long as I can remember, along with 29 discoveries in marketing, scientific methodology, social responsibility, strategic planning, education, and applied statistics. They are my legacy. And I hope to discover more. In addition to continuing to work on global warming and political forecasting, I am involved in efforts to improve advertising—an area

of the economy which is quite inefficient. This represents a converging of my interests in persuasion and the index method.

Finally, I thank those editors who have been inviting me to publish papers. Judging from life-expectancy forecasts, I should avoid having to spend five years in the reviewing process.

References:

Armstrong, J. S. (1980). The seer-sucker theory: the value of experts in forecasting. Technology Review, 83, 18-24.

Armstrong, J. S. (1982). Research on scientific journals: Implications for editors and authors. Journal of Forecasting, 1, 83-104.

Armstrong, J. S. (1983). Cheating in management science. Interfaces, 13, 20-29.

Armstrong, J. S. (1986). Research on forecasting: A quarter-century review, 1960-1984 (with commentary). Interfaces, 16, 89-109.

Armstrong, J. S. (1997). Peer review for journals: evidence on quality control, fairness, and innovation. Science and Engineering Ethics, 3, 63-84.

Armstrong, J. S. (2011). Natural learning in higher education. In Encyclopedia of the Sciences of Learning. Heidelberg : Springer.

Armstrong, J.S. (2012), Illusions in regression analysis. International Journal of Forecasting (forthcoming).

Armstrong, J. S., & Pagell, R. (2003). Reaping benefits from management research: Lessons from the forecasting principles project. Interfaces, 33(6), 89-111.

Chamberlin, T. C. (1965). The method of multiple working hypotheses. Science, 148, 754-759. (Reprint of an 1890 paper.)

Collopy, F., & Armstrong, J. S. (1992). Rule-based forecasting: development and validation of an expert systems approach to combining time series extrapolations. Management Science, 38, 1394-1414.

Evanschitzky, Heiner & J.S. Armstrong (2010), Replications of research on forecasting,. International Journal of Forecasting, 26 (2010), 4-8.

Fildes, R., & Makridakis, S. (1995). The impact of empirical accuracy papers on time series analysis and forecasting. International Statistical Review, 63, 289-308.

Frey, B. S. (2003). Publishing as prostitution? – Choosing between one's own ideas and academic success. Public Choice, 116, 205-223.

Gardner, E.S. and McKenzie, E. (1985) Forecasting trends in time series. Management Science, 31(10), 1247-1246.

Graefe, A., Armstrong, J. S., Jones, R., & Cuzán, A. G. (2012). Combining forecasts. Working paper.

Green, K. C., & Armstrong, J. S. (2007). Global warming: Forecasts by scientists versus scientific forecasts. Energy and Environment, 18(7-8), 995-1019.

Green, K. C., & Armstrong, J. S. (2011). Effects of the global warming alarm: A forecasting project using the structured analogies method. Working Paper.

Green, K. C. & Armstrong, J.S. (2012), Commercial free speech: Evidence on the effects of mandatory disclaimers. Working Paper: http://kestencgreen.com/g&a-mandatory.pdf

Kealey, T. (1996). The economic laws of scientific research. London: Macmillan.

Lawrence, M. J., Edmundson R. H., & O'Connor M. J. (1985). An examination of the accuracy of judgmental extrapolation of time series. International Journal of Forecasting, 1, 25-35.

Schroter, Sara, Nick Black, Stephen Evans, et al. (2008). What errors do peer reviewers detect, and does training improve their ability to detect them? Journal of the Royal Society of Medicine, 101, 507–514.

Tetlock, P. E. (2005). Expert political judgment: how good is it? How can we know? Princeton University Press.

Wright, M., & Armstrong, J. S. (2008). Verification of citations: Fawlty Towers of knowledge. Interfaces, 38(2), 125-139.