SOCIETY OF ACTUARIES   Forecasting and Futurism
Section

ISSUE 2 | JULY 2010

Forecasting & Futurism
NEWSLETTER

Best Methods and Practices in Judgmental Forecasting

By Alan Mills

Actuaries
Risk is Opportunity.®

**2009-2010
SECTION LEADERSHIP**

Chairperson
Alan Mills
FSA, ND
Alan.Mills@earthlink.net

Vice Chairperson
Ben Wadsley
FSA, MAAA
bwadsley@Aegonusa.com

Secretary/Treasurer
Scott McInturff
FSA, MAAA
Scott.McInturff@sunlife.com

COUNCIL MEMBERS

Philip Bieluch
FSA, MAAA, FCA
Life Meeting Liaison
phil@bieluch.com

Min Deng
ASA, Ph.D.
Co-Education Liaison
mdeng@maryville.edu

Jon Deuchler
ASA, MAAA
Co-Research Coordinator
Jdeuchler@rgare.com

Michael Lindstrom
ASA, MAAA
Research Coordinator
mlindstrom@federallife.com

David Snell
ASA, MAAA
Relationship Coordinator
dsnell@rgare.com

Raza Zaidi
ASA, MAAA
Web Liaison
Raza.Zaidi@lfg.com

BOARD PARTNER

Philip Gold
FSA, MAAA, FIA
Phil.Gold@ggy.com

NEWSLETTER EDITOR

David Snell
ASA, MAAA
dsnell@rgare.com

SOA STAFF

Jacque Kirkwood, Staff Editor
jkirkwood@soa.org

Meg Weber, Staff Partner
mweber@soa.org

Christy Cook, Section Specialist
ccook@soa.org

Julissa Sweeney, Graphic Designer
jsweeney@soa.org

Judgmental Methods, Collaboration, Contests and More!

By Dave Snell

Welcome! We are beginning a series of themed newsletter issues—each focused on a particular type of forecasting. Our theme for this issue is judgmental methods. I have to admit that the lead article, "Best Methods and Practices in Judgmental Forecasting," by Alan Mills, is a bit disconcerting to me. As usual, Alan's writings enlighten and educate; but it is a disappointing wake-up call to discover that the tried and true methods most commonly used by actuaries are certainly popular and often tried, but often not true. Expert opinions, intuition and traditional meetings are exposed as generally less accurate than simple quantitative methods. However, there are some judgmental methods, such as Delphi studies, that seem to do well; and Ben Wolzenski's article recapping the SOA Blue Ocean study "Future Opportunities in the Life Insurance Industry—Views of a Delphi Study" tells us about one such collaborative success. Scott McInturff adds an unconventional and intriguing review of *The Wisdom of Crowds*, a book that further explains the power of collaboration.

In the spirit of collaboration, our section has been reaching out to network with other SOA sections and with external organizations such as the World Future Society (WFS) and the Santa Fe Institute (SFI) – the birthplace for the sciences of complexity. "A New End, A New Beginning" by John Petersen, is reprinted from the September October 2009 issue of *The*

Futurist (the WFS magazine); and I have included a review I wrote of *Complexity: A Guided Tour,* by Melanie Mitchell, an external professor at the SFI. Regarding our cooperation with other SOA sections, I'm proud to announce that the SOA Annual Meeting this year will see complexity science sessions cosponsored by the Forecasting & Futurism, Actuary of the Future, and Health Sections; and we have council member ties now to these sections plus the Technology and Investment Sections. If you are active with another section and would be willing to expand our network of kindred spirits, please contact any of us. We'd like to meet you.

The joint session we cosponsored with The Financial Reporting and Investment Sections to bring Nassim Taleb to last year's annual meeting is summarized by Ben Wadsley in his article, "Living with Actuarial Black Swans—a Discussion with Nassim Nicholas Taleb." This event drew a very large attendance, and in another planned session for the 2010 Annual Meeting in New York, we shall address Taleb's question about whether prediction is even possible.

A lot is happening on the forecasting and futurism front, and this is an exciting time to be a part of it. Alan's chairperson's message announces a contest we hope many of you and your colleagues will enter to win fame and an Apple iPad and

further advance the state of forecasting. Granted, it's not a $1 million prize like the one NetFlix offered (see the article, "Forecasting Judgment: The Netflix Prize and Collaborative Filtering," by Mike Lindstrom); but each of us can help the profession with our contributions.

To paraphrase Pogo in the famous cartoon from Earth Day, 1971, "We have met the *future*, and *she* is us." Enjoy the articles in this issue. Think about your article for the next issue; the theme is going to be simulation methods. Become an active participant in our future. ▼

Cheers,

Dave

Environmental Notice!

No trees were harmed in the distribution of this newsletter. It is our first issue using electronic media. However, we realize that some members do not wish to read large amounts of information on a computer monitor; and others may wish to have it available offline. One option, of course, is to print it out on your printer; but that defeats some of the ecological intent here. Instead, please consider downloading it to your Kindle, Nook, iPad or other electronic reader. Both the Kindle (from Amazon) and the Nook (from Barnes and Noble) have the capability of displaying a PDF.

However, neither the Nook nor the Kindle do a perfect job of rendering PDF documents. If your eyes are able to read small print, the Kindle can show an accurate copy of the PDF if you just copy it over to the documents folder. The Nook has built-in software to allow font scaling for easier readability; but it sometimes changes formatting. The Kindle also offers the scalable fonts via your Kindle user e-mail account. This can be accomplished by sending the PDF document as an e-mail attachment to your Kindle e-mail account (e.g., *mykindlename@free.kindle.com*) with the word 'convert' in the subject heading. This will result in an e-mail reply (usually within a few minutes) that has a file for you to copy to your Kindle. If you send to the non free account (e.g., *mykindlename@kindle.com*) then Amazon will send it wirelessly to your Kindle for you, at a cost of 10 cents.

Both of these readers allow you to take your newsletter along with you. Quoting the late Karl Malden, "Don't leave home without it!"

Dave Snell

Dave Snell, ASA, MAAA, is technology evangelist with RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *dsnell@rgare.com*.

Want to Win an iPad?

By Alan Mills

This issue initiates a new feature: a **forecasting application contest**. The Forecasting and Futurism Section will award a **16GB iPad** to the section member who develops the best actuarial application of a judgmental forecasting method, incorporating best practices outlined in this issue's article titled "Best Practices in Judgmental Forecasting." In addition, we will feature the winning application in an issue of the newsletter.

To enter the contest, apply one or more judgmental forecasting methods to forecast an event or quantity applicable to actuarial work, and write a brief essay describing your method, how you implemented it, its results (if any), and how it compares to prior methods. Send your entry to Christy Cook (at *ccook@soa.org*) before Dec. 1, 2010. Only members of the Forecasting and Futurism Section are eligible to enter this contest. For contest details, visit the Latest News section of the Forecasting and Futurism Web page: *www.soa.org/forecasting-futurism.*

CURRENT ACTIVITIES

You can be proud of your section council members and friends; they are working hard. Following are some of our current projects:

• **Newsletter.** In 2010, we will publish two newsletters, the one you're reading and another in the fourth quarter. The theme of the fall newsletter will be simulation modeling, with an emphasis on agent-based modeling and complexity science.
• **Forecasting competitions.** In addition to the application contest described above, this year we will also sponsor a major forecasting competition for all actuaries. We will announce the competition at the SOA Annual Meeting in October.
• **Meeting sessions.** This year we are presenting many sessions at actuarial meetings. At the Life and Annuity Spring Meeting, we sponsored two sessions titled, "Predictable or Non-Predictable: Forecasting in the 21st Century" and "Product Development – Can We Admit Our Mistakes?" At the Health Spring Meeting we presented a session titled, "Applications of Complexity Science for Health Actuaries." And we will present two sessions at the SOA Annual Meeting, one titled "Is Prediction Possible?" and another about complexity science.
• **Survey.** We will soon survey you and other actuaries to find out how you use forecasting and futurism methods, and what you would like to learn about such methods. I urge you to provide your feedback; it will benefit all actuaries.
• **Wiki.** We are exploring Wiki technology in order to share information about forecasting and futurism methods with actuaries worldwide. The planned Wiki will provide information about forecasting and futurism tools that are useful for actuaries. It will also have a discussion forum for members.

Even though you have a vigorous section council, **we need your help.** If you would like to participate in any of these activities, or have suggestions about other projects you'd like us to consider, we warmly welcome you and your ideas. Please contact me (at *alan.mills@earthlink.net*) or any council member. E-mail addresses can be found on the inside front cover of this issue.

We hope you enjoy this issue. (Let us know what you think.) ▼

Alan Mills

Alan Mills

Alan Mills, FSA, ND, is a family practice physician. He can be reached at *alan.mills@earthlink.net*.

Best Methods and Practices in
Judgmental Forecasting

By Alan Mills

"We should be quite careful in trusting the intuitions of experts."

Daniel Kahneman[1]

All forecasting is, at least in part, judgmental forecasting. No matter how formal, computerized or technically sophisticated your forecasting method, it is rooted in human judgment. Judgment is required to select the method, to choose its parameters, to filter its data and to interpret its results.

And some forecasting, even in actuarial work, is mainly judgmental. Actuaries sometimes forecast results using mainly their experience, "gut feel," opinions, and intuition; and chief actuaries often use their judgment to summarily override all model results. In the business world, such behavior is common: A 2003 survey of 240 U.S. corporations found that only 11 percent use formal nonjudgmental forecasting methods at all, and of these 60 percent routinely adjusted their forecasts based on judgment.[2] Economists are no different: judgment is "the primary factor that the economist uses in converting mere statistical and theoretical techniques into a usable forecast."[3]

Given the ubiquity of judgmental forecasting, it is important to understand the potential biases inherent in human judgment—the traps that can snare us—and to learn methods and best practices to avoid them. The aim of this article is to help you with these goals.

BIAS AND ERROR IN HUMAN JUDGMENT

Our understanding of bias and error in human judgment comes mainly from the work of Daniel Kahneman and Amos Twersky *(see sidebar, Kahneman and Twersky)*. Their experiments in human judgment and decision-making, together with the experiments of scientists who followed them, uncovered a startling result: much of human judgment is based not on rational cognitive processes, but

Alan Mills

Alan Mills, FSA, ND, is a family practice physician. He can be reached at *alan.mills@earthlink.net.*

Kahnemann and Twersky

Psychologists Daniel Kahneman and Amos Twersky shared one of the most productive collaborations in the history of social science. Starting in 1969, for more than 25 years they conducted groundbreaking experimental research into human judgment and decision-making. Their research had such a profound impact that in 2002 Kahneman became the first psychologist to win a Nobel Prize in economics (an honor that, had he lived, Tversky would have shared).

As an example of one of their experiments: In two trials, participants immersed a hand in cold water until instructed to remove it. The first trial lasted 60 seconds at 57 degrees Fahrenheit (very painful), and the second trial lasted a total of 90 seconds with 60 seconds again at 57 degrees followed by 30 seconds at 59 degrees (a little less painful). When asked which of the two trials they would choose to repeat, the remarkable finding is that 65-80 percent of subjects elected to repeat the second trial, even though it was longer than the first trial and produced more pain.

This, and a host of similar experiments, led Kahneman and Twersky to conclude that we store memories of our experiences according to what they called a "peak/end rule heuristic:" our memory of events is primarily an amalgam of the peak point of the experience and its end point. Nothing else matters.

For an excellent introduction to their work, see the YouTube videos Kahneman (2008a) and Kahneman (2008b).

FOOTNOTES
[1] Kahneman (2008a)
[2] Sanders & Manrodt (2003)
[3] McAuley, 1986, p. 384

rather on *heuristics*—unconscious "rules of thumb" that humans have developed over millennia to deal with our environment. And, although our ingrained and unconscious heuristics may have served us well in dealing with the dangers and opportunities faced by hunters and gatherers, in today's complex world, they can produce serious errors.

Following are research-based examples of heuristics, biases (i.e., heuristics that are systematically skewed from rationality), and common cognitive difficulties that often cause us trouble when we make judgmental forecasts:[4]

Peak/end rule heuristic: In an experience, we tend to remember only the most extreme point, and the final point (the peak and the end). Thus, our memories—even the memories of experts are highly inaccurate. Since our intuitions are largely based on memory, our intuitions also can be highly inaccurate.

Framing bias: We judge an issue according to how it is presented. Many times Kahneman and Twersky demonstrated that opinions about an issue can be reversed if the issue is simply presented in a way that is logically equivalent, but expressed differently.

Anchoring bias: In our judgments, we often rely too heavily on one piece of information. For example, in one of their first studies, Kahneman and Twersky asked people to write down the first three numbers of their telephone number, and add 400 to it. They asked the subjects to consider this number as a year AD, and then asked them to guess when Attila the Hun was defeated in Europe. Invariably, the resulting guess was very close to the result of the addition.

Representativeness bias: We tend to make judgments based on small samples that are not statistically representative.

Availability bias: We make judgments based on data that is easily available, rather than finding appropriate data.

Confirmation bias: We focus on aspects of the past that conform to our views, and generalize from these to the future. We are blind to what would refute our views, and

only look for corroboration. This is the central problem of induction: we generalize when we should not.

Conjunction bias: In an experiment, researchers found that when a terrorist attack had occurred recently, people about to board a plane are willing to pay more for insurance that covers terrorism than for insurance that covers any cause of death, including terrorism. Thus, recent significant events cloud our ability to reason.

Narrative bias: We automatically fabricate stories, weaving narrative explanation into a sequence of historical facts, and thereby deceive ourselves that we understand historical causes and effects and can apply this understanding to the future. This bias gives us a false sense of forecasting confidence, a sense that the world is less random and complex than it really is—a complacency that leads to forecast error.

Proximate cause bias: In our search for cause and effect relationships, we tend to consider only the most proximate causes.

Expert bias: We overvalue expert opinion.

Difficulty judging probabilities: Kahneman and Tversky found that people, even those who are statistically sophisticated, are not good at judging probabilities:

- When people are asked to estimate the probability that a randomly selected group of men has an average height over six feet, they give about the same probability whether the group consists of 10 men or 1,000.
- When presented with new information, people tend to ignore other probabilities. For example, suppose a reliable test for a rare medical condition is positive for

FOOTNOTES
4   Most of these can be found in Kahneman, Slovic, & Tversky (1982) and Kahneman & Tversky (2000).

you. The doctor tells you that one person in 10,000 has the condition, that for a person with the condition the test returns an accurate positive result 99 percent of the time, and that the test gives an accurate negative result 99 percent of the time for people who do not have the condition. Should you worry? Most people worry, because they focus on the new information, the positive test result, rather than the problem as a whole. They conclude that the chances are overwhelming that they have the disease, when in fact the chance is only about 1 in 100.

To understand this result, consider a group of 10,001 people who are tested for the disease. On average, only one of the people in this group actually has the disease, and 10,000 do not. For that one person, 99 percent of the time, the test will return a positive result. However, for the remaining 10,000 people, the test will return a positive result for 100 people (because the test returns an accurate negative result for only 0.99 x 10,000 = 9,900 people). Thus, if you are someone who received a positive result, you may be the one person who has the disease, or you may be one of the 100 for whom the test returned an incorrect positive result. Your chance of

having the disease is therefore 1/101, or about 1 percent. You need not worry too much.

Overconfidence: When asked for the probability that their prediction of some event will come true, people—especially experts—systematically report a probability that is far too high. Similarly, they are overconfident that particular disasters will not happen. As a dramatic example, using their judgment, NASA managers assessed the probability of failure for the space shuttle Challenger as 0.00001, even though their engineers assessed the probability as 0.01.[6] In studies performed by numerous researchers, experts provide judgmental prediction intervals for their forecasts that are far too narrow; they are overconfident in their forecasts by a wide margin, and especially so when in addition to prediction intervals they are asked to provide point predictions.[7]

The list of biases and cognitive difficulties goes on and on. But you get the point: no matter how much mathematics and statistics we study, or how much experience we have, our judgment is largely governed by unconscious heuristics, and is prone to substantial error. We are human and we err. (*see sidebar, Downright Humiliating for Experts*).

METHODS

To guard against our inherent biases that lead to judgmental forecasting error, there is much that we can do. First, we can follow methods that, according to research, are more accurate. The following chart on page 9 shows 10 common judgmental forecasting methods.

In the chart, the methods are arranged according to the type of forecasters (whether individuals, groups, individuals or group, or automated) and according to their degree of struc-

Downright Humiliating for Experts
"… one of the great classics in the history of psychology, work by Paul Miel, was a seminal study looking at all the previous studies that had compared the performance of experts (like clinical psychologists) predicting various criteria, to very simple linear combinations of variables. So, you have a clinical psychologist looking at a lot of information, with a subset of that information used in a statistical model, and you compare how well the intuitions of people do compared to the statistical model. The statistical model is based on only part of the information, and is applied in a restricted way as a simple linear combination of variables. The stunning result was that in all the studies that Miel looked at, the linear equations beat the experts hands down … a simple model does better.

Now, 50 years later, at last count there are 180 studies along the same lines, and basically the conclusion is the same: When you compare people to very simple combinations of variables, the combination wins just about every time. It's hard to find any exception. The results of this research are downright humiliating for people who try to forecast complicated events."
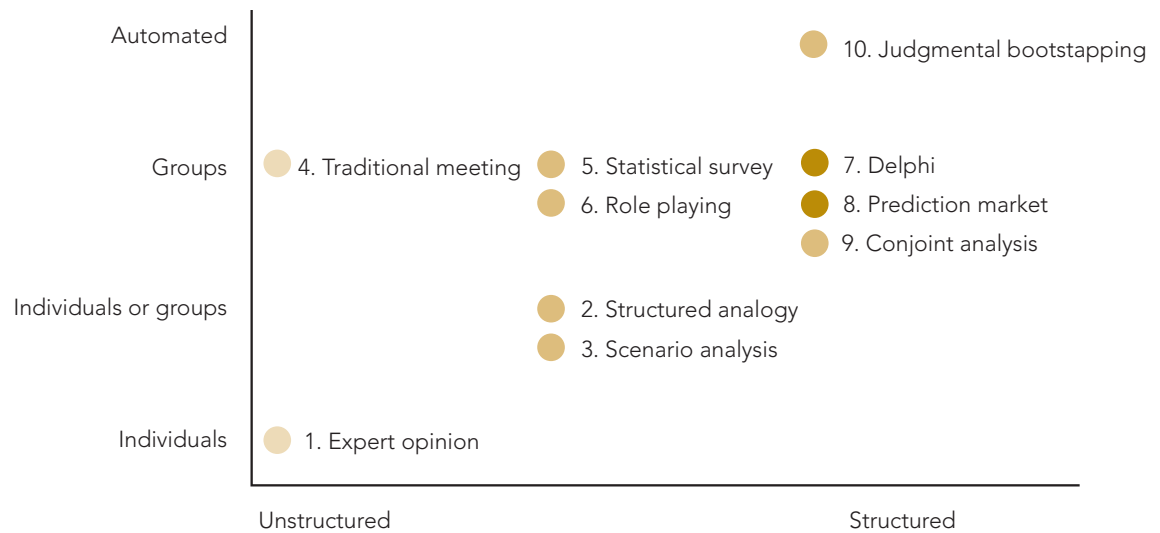Daniel Kahneman[5]

FOOTNOTES
[5] Kahneman (2008a).
[6] From the Feynman appendix (Appendix F) to the Rogers commission report on the space shuttle Challenger accident (*science.ksc.nasa.gov/shuttle/missions/51-l/docs/rogers-commission/table-of-contents.html*).
[7] Lawrence, Goodwin, O'Connor, & Onkal (2006), pages 505-506.

ture. Structured methods consist of systematic and detailed steps that can be described and replicated. The color depth of the dots in the chart represents the methods' relative accuracy; the deeper color is more accurate. As the chart shows, unstructured methods are generally less accurate than structured, and the individual method is less accurate than group methods.

Following are brief descriptions of these methods, in the order indicated on the chart.

Automated
                                                                    ● 10. Judgmental bootstapping

Groups   ● 4. Traditional meeting   ● 5. Statistical survey   ● 7. Delphi
                                        ● 6. Role playing   ● 8. Prediction market
                                                                   ● 9. Conjoint analysis

Individuals or groups   ● 2. Structured analogy
                                   ● 3. Scenario analysis

Individuals   ● 1. Expert opinion

Unstructured   Structured

Individual method

1. Expert opinion

Perhaps the most common judgmental forecasting method is to ask the opinion of an expert. Although common, this method is perhaps the most error-prone. It is generally unstructured, and fraught with all the biases and cognitive difficulties of an individual human.

Individual or group methods

2. Structured analogy

This method compares a recent series of events to a similar series that occurred earlier in another context. Forecasted outcomes are then based on past actual outcomes in the other context. The two series generally share important characteristics such as time scale, severity, reversibility, impacted sector, aggravating factors, etc. The purpose of the method is to constrain judgmental forecasting to a situation that actually happened in the past.[8]

3. Scenario analysis

The scenario analysis method is a process of forecasting future events by framing alternative possible outcomes in terms of story-like narrative scripts that often include the impact of events such as new technology, population shifts or changing consumer preferences. Usually, the method includes development of a most likely scenario, along with at least one optimistic and one pessimistic scenario. The primary purpose of a scenario is to constrain judgmental forecasting to a narrative structure, with the aim of producing more realistic forecasts. Scenario analysis is used by many organizations for long-term forecasting, and has been found to be quite accurate when used properly.[9]

Group methods

4. Traditional meeting

The most common method to obtain a judgmental forecast from a group of people is the traditional meeting, with unstructured discussion around a table. But this method has a number of drawbacks: The outcome is often unduly influenced by expert bias, difficulties of communication, psychological factors such as yielding to the opinions of

FOOTNOTES
[8] Armstrong (2001), Chapter 7: Analogies, pages 193-213.
[9] Mills & Bishop (2000)

5. Statistical survey

The statistical survey is a method of obtaining opinions about a specific topic from a sample of a population that is considered to be representative of the whole population. Statistical surveys are a widely used judgmental forecasting method.[10] For example, the Library Research Service recently surveyed libraries about the impact of e-readers on the future of libraries, and specifically about whether libraries will cease to exist. Used properly, a survey can produce a useful judgmental forecast.[11]

6. Role playing

In this method, people play roles to enact a situation in a realistic manner. It is particularly useful in forecasting the outcome of competition or conflict. The method has been shown to be substantially more accurate than expert opinion.[12]

RAND and Delphi

Olaf Helmer and Norman Dalkey of the RAND Corporation developed the Delphi method in 1953. It was based on a prior method, also developed by RAND, to combine the opinions of horse-racing handicappers to improve the chances of winning horse races.

Helmer and Dalkey thought a group of experts is the best way to develop a judgmental forecast (as Dalkey put it, "n heads are better than one"), but they knew from experience that a group of experts around one table can lead to argument and little progress. They designed Delphi to maximize the information that can be obtained from such a group.

In 1964, RAND published the first Delphi study, titled "Report on a long-range forecast." The study included a panel of 82 experts, including Issac Asimov, Arthur Clarke, Bertrand de Jouvenel, and Dennis Gabor. Its purpose was to forecast scientific and technological advances through the year 2000 and beyond. Most of the forecasts turned out to be amazingly accurate.[13]

(continued) authority figures or aggressive personalities, and the excessive influence of majority opinion.

7. Delphi

The Delphi method is a structured group consensus method for obtaining judgmental forecasts from experts *(see sidebar RAND and Delphi)*. It polls the experts anonymously, over successive rounds, with summary feedback in between. Over the course of a Delphi study, participating experts see where they stand in the group and may adjust their views accordingly. The feedback in successive rounds includes reasons for the more extreme views. Sometimes those reasons convince others that they are unwittingly making an erroneous assumption or ignoring an important piece of evidence. The result is a deeper exploration of the reasons behind expert opinions, without the biases, psychological impediments, and conflict that often appear in unstructured group discussions.

The Delphi method generally produces a rapid narrowing of opinions, and has been demonstrated to provide more accurate forecasts than unstructured group discussions. Interestingly, a face-to-face group discussion following the application of the Delphi method generally degrades forecast accuracy.[14]

8. Prediction market

A prediction market is another structured group consensus method, one that develops judgmental forecasts based on the mechanism of a speculative market. In this method, participants buy or sell shares of 'claims' regarding a particular forecast (e.g., the next president, an Oscar winner, or the increase in health care expenditures next quarter). If a claim turns out to be true, then one share is worth a stated amount (e.g., a claim is worth $1 if a particular candidate becomes president). A participant places a bet on the outcome by buying or selling shares at a market-determined price. For

---

FOOTNOTES

[10] Fowler (2009)
[11] Surowiecki (2004) and (2007).
[12] Armstrong (2001), Chapter 2, pages 13-30.
[13] Mills & Bishop (2000)
[14] Mills & Bishop (2000) and Adler & Ziglio (1996).

example, a participant might buy 100 'yes' shares that a certain candidate will become president at 60 cents per share. The participant would then win $100 if the candidate won. The current market price can thus be interpreted as the participants' forecast of the probability of the event or the expected value of the parameter. Prediction markets are strikingly prescient.[15]

9. Conjoint analysis

This method quantifies respondent judgments and opinions, by having the respondents trade conflicting event or object attributes against one another. Analysis of these trade-offs reveals the relative importance of the attributes. The method is often used to forecast consumer response to new products. For example, pharmaceutical companies use this method to understand physician opinions about drugs, in order to forecast product sales and market share.[16]

AUTOMATED METHOD

10. Judgmental bootstrapping

Judgmental bootstrapping is an automated expert system that models an expert's reasoning process. To develop the model, an expert's forecasts are regressed against the information the expert used to make the forecasts. Because such models apply the reasoning processes of experts in a consistent way, studies from psychology, education, personnel, marketing and finance have shown that bootstrapping forecasts are more accurate than forecasts made by experts with unaided judgment.[17]

BEST PRACTICES

In addition to using methods that researchers recommend, we can employ well-researched best practices in applying these methods. Following is a review of such practices. They are organized according to the methods to which they apply.

Practices that apply to all methods

Provide feedback

One of the key findings of researchers is that records should be kept about judgmental forecasts, in order to provide the forecasters with feedback. Feedback is valuable because it enables the forecaster to learn. There are three main types of feedback:

- **Outcome:** The most common type, providing the forecaster with the latest observation in a series.
- **Performance:** Describes the accuracy and biases of the forecaster's forecasts.
- **Cognitive process:** Describes the strategy the forecaster used to arrive at the forecast. For example, such feedback might include a graphical display of the weights the forecaster attached to different data.

Researchers have found that outcome feedback is the least effective form.[18] For actuarial judgmental forecasts, a combination of all the feedback forms would likely be most useful.

Provide checklists

Provide the judgmental forecaster with a checklist of information categories relevant to the forecasting task. Checklists remind forecasters about factors relevant to their forecasts, and prevent them from being influenced by extraneous information.[19]

Present data clearly, in both graphs and tables

Present information to the forecaster clearly; in particular, avoid presentations that require forecasters to recognize complex patterns or to mentally aggregate many numbers. Because people vary in their ability to extract information from graphs and tables, present data in both formats.[20]

Frame questions in different ways

To avoid the framing bias, pose questions in various ways, from various perspectives.[21]

FOOTNOTES

[15] Servan-Schreiber, Wolfers, Pennock, & Galebach (2004) and Surowiecki (2004).

[16] Armstrong (2001), Chapter 5: Conjoint analysis, pages 145-167.

[17] Armstrong (2001), Chapter 6: Judgmental bootstrapping, pages 169-192.

[18] Lawrence, et al. (2006), page 507; and Armstrong (2001), page 63.

[19] Armstrong (2001), page 61.

[20] Lawrence, et al. (2006), page 497-498; and Armstrong (2001), page 64 and 93.

[21] Armstrong (2001), page 697.

Use mechanical methods

To help forecasters process complex information, especially statistical information, use mechanical methods rather than relying on judgment or mental processes.

Combine forecasts

Researchers have found that combining judgmental forecasts with either statistical forecasts or with other judgmental forecasts improves forecast accuracy.[22]

Practices that apply to groups

Use heterogeneous groups

A forecast developed by a group, especially a heterogeneous group, is generally more accurate than one by an individual, even if the individual is an expert. Generally, the various structured consensus methods will produce more accurate results than an individual expert.[23]

Employ an adequate number of forecasters

In surveys and prediction market models, make sure that the sample size is adequate to represent the entire population. In expert consensus methods, use between five and 20 experts.[24]

Pretest questions

Prior to data collection, questions should be tested on a sample of potential forecasters to ensure that they are understood and that they relate to the objectives of the problem.

Practices that apply to groups of experts

Use heterogeneous experts

If you use an expert consensus method, such as the Delphi method, make sure the experts are heterogeneous, that they vary in their information sources and in the way they approach the problem.

Request justification in writing

Experts should provide the reasons for their forecast, in writing. The Delphi method incorporates this best practice.

Provide numerical scales with several categories

To avoid anchoring bias, use as many categories of potential forecast responses as reasonable.[27]

Require multiple forecasts

Ask experts to make forecasts, and then repeat the process some days later. The Delphi method incorporates this practice.[28]

Require confidence intervals

Require experts to use confidence intervals, rather than point predictions.

Early in his career, Daniel Kahneman coined the term "illusion of validity," to capture the truth that we—especially if we are "experts"—often harbor an illusion that we are good at judgmental forecasting, when in fact we are not. In his research, he and Amos Twersky showed us that the accuracy of our judgment is severely compromised by inherent and unconscious cognitive weaknesses. Based on their work, others have developed methods and best practices to circumvent these weaknesses, and improve our judgmental forecasts. In your work and the work of your colleagues, do you see any opportunities to use these methods and best practices to improve your judgmental forecasts?

FOOTNOTES
[22] Lawrence, et al. (2006), page 508.
[23] Surowiecki (2004), (2007), and Armstrong (2001) page 698.
[24] Armstrong (2001), pages 698-699.
[25] Armstrong (2001), page 698.
[26] Armstrong (2001), page 697.
[27] Armstrong (2001), page 698.
[28] Armstrong (2001), page 699.

REFERENCES

Adler, M., & Ziglio, E. (1996). *Gazing Into The Oracle: The Delphi Method and Its Application to Social Policy and Public Health.* London: Jessica Kingsley Publishers, xi, 252 p.

Armstrong, J. S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners.* Boston, Mass.: Kluwer Academic, xii, 849 p.

Fowler, F. J. (2009). *Survey Research Methods* (4th ed.). Thousand Oaks: Sage Publications, x, 201 p.

Kahneman, D. (2008a). "Explorations of the Mind — Intuition: The Marvels and the Flaws," *Hitchcock Lectures*. Retrieved from *www.youtube.com*. 56 minutes.

Kahneman, D. (2008b). "Explorations of the Mind — Well-being, *Hitchcock Lectures.* Retrieved from w*ww.youtube.com*. 59 minutes.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment Under Uncertainty:* Heuristics and Biases. Cambridge; New York: Cambridge University Press, 555 pages.

Kahneman, D., & Tversky, A. (2000). *Choices, Values, and Frames.* New York; Cambridge, UK: Russell sage Foundation; Cambridge University Press, 840 pages.

Lawrence, M., Goodwin, P., O'Connor, M., & Onkal, D. (2006). "Judgemental Forecasting: A Review of Progress Over the Last 25 Years. *International Journal of Forecasting,* 22, 493-518.

Mills, A., & Bishop, P. (2000). *"Applied Futurism: An Introduction for Actuaries.* Society of Actuaries, 128 pages.

Page, S. E. (2007). The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton: Princeton University Press, xxiv, p. 424 .

Sanders, N. R., & Manrodt, K. B. (2003). "The Efficacy of Using Judgmental Versus Quantitative Forecasting Methods in Practice. *Omega* (31), 511-522.
Servan-Schreiber, E., Wolfers, J., Pennock, D. M., & Galebach, B. (2004). "Prediction Markets: Does Money Matter? *Electronic Markets,* 14 (3 - September 2004), 243-251.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations* (1st ed.). New York: Doubleday, xxi, 296 p. ▼

Future Opportunities in the Life Insurance Industry—Views of a Delphi Study

By Ben Wolzenski

The following article contains numerous excerpts or paraphrases from the February 2009 SOA study, "Blue Ocean Strategies in Technology for Business Acquisition by the Life Insurance Industry," the full text of which is available on the SOA website under Research of the Forecasting & Futurism Section.

INTRODUCTION

In 2007, members of the Forecasting & Futurism Section were joined by members of the Technology Section and the Marketing & Distribution Sections to conduct a Delphi study (*see sidebar, "What's a Delphi Study?"*) about possible developments in the life insurance industry. We named it "Blue Ocean Strategies in Technology for Business Acquisition by the Life Insurance Industry." Our goal was to have a panel of experts—actuaries and other financial professionals from across the country—identify and debate possible new approaches to acquiring business by life insurers. Panelists were asked to answer questions about possible "Blue Ocean Strategies" *(see sidebar, "What's a Blue Ocean Strategy?")*. Three rounds of surveys were performed over a period of 14 months, leading to what our report calls "a creative compendium of insightful ideas on the subject."

What's a Delphi Study?
According to Wikipedia:
"The Delphi method is a systematic, interactive forecasting method which relies on a panel of independent experts. The carefully selected experts answer questionnaires in two or more rounds. After each round, a facilitator provides an anonymous summary of the experts' forecasts from the previous round as well as the reasons they provided for their judgments. Thus, participants are encouraged to revise their earlier answers in light of the replies of other members of the group … the process is stopped after a predefined stop criterion (*e g.*, number of rounds, achievement of consensus, stability of results) and the mean or median scores of the final rounds determine the results."

By its very definition, a Blue Ocean Strategy is not likely to generate anything even approaching a mean or median score, which makes this a somewhat unusual topic to be subjected to a Delphi Study. In this study panelists retained very different views about which scenarios will play out, and much can be learned from the reasons given for those differing views.

What's a Blue Ocean Strategy?
A Blue Ocean Strategy has come to mean a strategy that defines and capitalizes on a vast open area of previously undiscovered and unoccupied territory. The concept is explored in the book, *Blue Ocean Strategy*, by W. Chan Kim and Renée Mauborgne, which was based on a decade-long study of 150 strategic moves spanning more than 30 industries over 100 years (1880-2000). Further details are available at the official website, *www.blueoceanstrategy.com*. Some examples of companies in the book:
- Cirque de Soleil: Blending of opera and ballet with circus format while eliminating star performer and animals;
- Netjets: fractional jet ownership;
- Curves: redefining market boundaries between health clubs and home exercise programs for women.

ROUND ONE—THE MIXING BOWL FOR THE PROJECT

The Round One Survey gave panelists background information on the Delphi Method and the Blue Ocean Strategy concept, then asked a series of 10 questions (*see sidebar, "Illustrative Round One Questions" about "possible current and future technologies and new strategies in business acquisition for the life insurance industry enabled by these technologies."*) Responses were received from 39

Ben Wolzenksi

Ben Wolzenksi, FSA, MAAA, is president of Actuarial Collaboration, Inc. He can be reached at *bwolzenski@rgare.com*.

of the 46 selected panelists. The open-ended nature of the questions produced a substantial volume and great variety of responses that covered over 70 pages. Given the great diversity of responses, the working group felt that the best approach to Round Two would be to synthesize the results into a finite number of "strategies." Ten was selected as a reasonable number of strategies, and the working group then constructed 10 "Blue Ocean Strategies" incorporating as much as possible from responses of all of the panelists.

Illustrative Round One Questions
- *Question No. 1.* What possible current or future technology could enable the life insurance industry to adopt a Blue Ocean Strategy *in how it markets its products* and what is the resulting Blue Ocean Strategy? Feel free to consider marketing methods currently employed by the insurance industry, that are currently employed by any other industry, or that, in your opinion, could and should be employed. Consider any existing or potential media.
- *Question No. 6.* What possible current or future technology could enable the life insurance industry to adopt a Blue Ocean Strategy *in how it processes applications for its products* and what is the resulting Blue Ocean Strategy?"
- *Question No. 10.* What emerging technologies do you see on the horizon with the potential to impact our daily lives? How could these impact the design, marketing, sales and/or processing of insurance?

ROUND TWO—THE INGREDIENTS INTO THE OVEN

The Round Two Survey strategies were presented to the participating panelists as follows: "This Round Two survey contains ten (10) strategies that represent a composite of many of the ideas contained in the Round One responses, and a series of questions about those strategies. Please complete any or all of the questions for which you have ideas; you do not need to respond to every question." The strategies were constructed from participants' responses made with the Round One instruction to "assume the technologies and strategies **could occur in the next 10 years.**"

Following are three of the composite strategies, along with, in each case, one of the questions about the strategy that the panelists were asked. How would you answer each question? Panelists' responses are shown at the end of this article.

STRATEGY NO. 1: EARTH FRIENDLY INSURANCE COMPANY—PAPERLESS PROCESSING

Earth Friendly Insurance Company plans to adopt a Blue Ocean Strategy called: "Paperless processing: do it all online!" Part 1 of this strategy is to use technologies and processes that do away with paper applications, which may include the prepopulation of some information about the applicant from internal or external sources. Information will be obtained through the Internet or all-in-one communication devices either directly from the applicant or a field agent. Policy approval and an option to print coverage verification will be directed back by similar routes. Earth Friendly also foresees a Part 2 of this strategy: the use of a "Touch the Screen" system in which the applicant would touch the computer/lap top screen and the finger print would automatically pull all medical files and other life style data. One slight prick of blood, similar to that used by diabetics for blood sugar testing, would provide immediate analysis of all physical conditions, which would be fed through the computer at the same time as the one-touch activity. One company has already adopted a version of Part 1 of this strategy, issuing up to $250,000 of term life coverage to individuals age 18 to 60 "generally within minutes" based on "just a few health questions" answered online. An immediate decision is provided and, if approved, the applicant can print their in-force policy online.

model as well as for each insurance category, which could be a broad spectrum (life, health, annuities, long-term care, auto and home) or some subset. Only products with relatively simple and transparent pricing would be offered. Consumers would mix and match discrete, simple products to address comparatively complex needs. Due to state insurance department restrictions, "Your Way" expects to issue multiple policies through different operating units to provide the overall coverage designed by the consumer. Online underwriting mechanisms and databases would be used to narrow the price range, define the price subject to certain conditions, or determine the price precisely. Response activity would be used to systematically refine the process model and coverage building blocks available to consumers.

Reader Response Opportunity—Strategy No. 1

Do you think "Part 2" of the strategy will become feasible in the next 10 years?

☐ A. Yes, achievable
☐ B. Probably achievable
☐ C. No, not achievable
☐ D. No opinion

See how the expert panel responded at the end of this article.

Reader Response Opportunity—Strategy No. 5

Is this a "Blue Ocean Strategy" or simply a window of opportunity for the early players?

☐ A. Blue Ocean Strategy
☐ B. Window of opportunity
☐ C. Not a Blue Ocean Strategy
☐ D. Neither
☐ E. No opinion

See how the expert panel responded at the end of this article.

STRATEGY NO. 5: YOUR WAY INSURANCE CO.—PROSPECTS CUSTOM-DESIGN COVERAGE ONLINE

A think tank at Your Way Insurance Company has recommended a "Blue Ocean" strategy in which individuals would custom-design their insurance coverage online. The entry point would be an online process driven model that enables consumers to design their insurance coverage by answering a series of questions. The model would have "click to call" expert advice available on how to use the

STRATEGY NO. 7: JUST WHAT YOU WANT INSURANCE COMPANY—MICROPOLICIES

Just What You Want Insurance Company believes that there may be an emerging opportunity for a Blue Ocean Strategy around offering micropolicies. These products cover narrow risks, at targeted periods, for specific consumers, at highly specialized prices. Sophisticated—often diverse—technologies are often required to enable distribution, segment markets, price risk and issue coverage. Although these policies have the potential to replace

broader blanket coverages, the greater potential is to open markets for risks otherwise uninsurable. For example, life insurance for a bungee jumper could be sold to cover the specific event.

Reader Response Opportunity—Strategy No. 7

Which of the following risks could be insured through a micro-policy? **(multiple answers permitted)**

☐ A. Animals (race horses, pets)
☐ B. Foreign travel activities in hazardous countries (vacation, humanitarian aid, journalists, etc.)
☐ C. Hazardous sports (bungee jumping, rock climbing, private plane piloting, space travel, etc.)
☐ D. Lifestyle choices (public speaking, extra-marital affair, political campaign)
☐ E. Specific diseases, medical conditions or procedures (HIV positive, heart surgery, LASIK, experimental surgery, etc.)
☐ F. Very specific event/risk (travel delay for executives, hostage, key witness, roadside bomb, specific body part for famous person, complications from surgery)

See how the expert panel responded at the end of this article.

ROUND THREE—THE ICING ON THE CAKE

In Round Three, panelists were provided with a summary of the Round Two responses, and asked if they had any additional input. Responses to Round Three were not extensive; the great majority of panelists did not indicate any further thoughts beyond their prior responses. This is one of the indicators that a Delphi Study is complete: the panelists are not changing their stances.

SUMMARY

Our goal was to collect new ideas; many of our Round One questions were open-ended to encourage creative responses. This worked well. In November of 2007, we received over 70 pages of summarized Round One responses from the SOA research staff. This was not a study where we could easily categorize the responses and apply meaningful statistical metrics to summarize them.

Our objective and expectation was not complete agreement among the respondents, but rather to surface new ideas; to vet them among a group of insurance, technology, and marketing and distribution professionals across the country; to distill them to a set of strategies deemed by some to have potential; and to elicit perceived obstacles and shortcomings of the strategies. In these respects, the project was successful.

The working group and SOA research staff members contributed more time than we originally planned because we found the interactions stimulating and thought provoking. The panel members gave a notable collective effort. They responded with enthusiasm to the questions as written—and to rewritings they occasionally added themselves. The anonymity condition seemed to remove inhibitions and open the tap for an outpouring—no, a deluge—of creative ideas.

Note: See how the expert panel responded on page 18. ▼

> ALTHOUGH THESE POLICIES HAVE THE POTENTIAL TO REPLACE BROADER BLANKET COVERAGES, THE GREATER POTENTIAL IS TO OPEN MARKETS FOR RISKS OTHERWISE UNINSURABLE. FOR EXAMPLE, LIFE INSURANCE FOR A BUNGEE JUMPER COULD BE SOLD TO COVER THE SPECIFIC EVENT.

EXPERT PANEL RESPONSES

Reader Response Opportunity—Strategy No. 1

Do you think Part 2 of the strategy will become feasible in the next 10 years?

- **Expert Panel Response**
  - ☐ A. Yes, achievable – 15 (54%)
  - ☐ B. Probably achievable – 6 (21%)
  - ☐ C. No, not achievable – 5 (18%)
  - ☐ D. no opinion – 2 (7%)

Reader Response Opportunity—Strategy No. 5

Is this a Blue Ocean Strategy or simply a window of opportunity for the early players?

- **Expert Panel Response**
  - ☐ A. Blue Ocean Strategy – 9 (32%)
  - ☐ B. Window of opportunity – 7 (25%)
  - ☐ C. Not a Blue Ocean Strategy – 3 (11%)
  - ☐ D. Not feasible, so neither – 1 (4%)
  - ☐ E. No answer – 5 (18%)
  - ☐ F. Wrote an expanded strategy to be Blue Ocean – 1 (4%)
  - ☐ G. Other comments – 2 (7%)

Reader Response Opportunity—Strategy No. 7

Which of the following risks could be insured through a micropolicy? **(multiple answers permitted)**

- **Expert Panel Response**
  - ☐ A. Animals (race horses, pets) – 2 (7%)
  - ☐ B. Foreign travel activities in hazardous countries (vacation, humanitarian aid, journalists, etc.) – 3 (11%)
  - ☐ C. Hazardous sports (bungee jumping, rock climbing, private plane piloting, space travel etc.) – 8 (29%)
  - ☐ D. Lifestyle choices (public speaking, extramarital affair, political campaign) – 3 (11%)
  - ☐ E. Specific diseases, medical conditions or procedures (HIV positive, heart surgery, LASIK, experimental surgery, etc.) – 7 (25%)
  - ☐ F. Specific event/risk (executive travel delay, hostage, key witness, roadside bomb, body part for famous person, complications from surgery) – 6 (21%)

The Wisdom of Crowds
by James Surowiecki

Reviewed by Scott McInturff
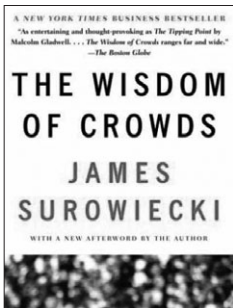
Trivial Pursuit is a game in which, in order to win, numerous questions related to general knowledge and pop culture must be correctly answered. Imagine a contest of Trivial Pursuit played by three distinct teams composed entirely of individuals with an actuarial background. Team One is composed of a single member who was the highest paid actuary in North America in the most recent calendar year. Team Two is a group of the five actuary wannabes who scored the highest on the most recent P/CAS Exam 1. Team Three is a group of 10 randomly chosen actuaries who are members of a North American Actuarial Society. Which team would you bet on to win the Trivial Pursuit contest?

TEAM ONE—INTELLECTUAL AND FINANCIAL SAVVY

Team One is likely comprised of a very intelligent actuary who is certainly the most financially successful participant in the game. Because of this intellectual and financial savvy, Team One might be expected to be able to out-think any other participant on any other team. It is quite probable that Team One is older than the average age of the other two teams participating in the contest, assuming that increasing financial success typically occurs with increasing age. Therefore Team One has additional years of experience to call upon in answering the general knowledge Trivial Pursuit questions. Based on the credentials of having the highest pedigree, Team One has an excellent chance to win the Trivial Pursuit game.

TEAM TWO—MASTERY OF PROBABILITY

Team Two is composed of five very bright individuals as evidenced by their superior mastery of probability relative to their peer exam takers. They are also likely to be five very young persons since they have recently completed one of the first actuarial exams. Their average age is most certainly younger than the average age of either of the other two teams. Their youth will likely give them some advantages on pop culture questions. Because of their youth, their life experience is lacking. However, Team Two may be able to compensate for this by having the collective experience of five members to draw upon in answering general knowledge questions. Based on their superior collective intelligence, Team Two will be a formidable competitor.

TEAM THREE—LIKELY TO LOSE OR MORE LIKELY TO WIN?

Team Three is somewhat of an enigma. Other than the fact that they are members of an actuarial society, we can't assign any details to any single member of the group to the extent we can with the first two teams. Because the 10 members were chosen randomly, it would be fair to assume that they are somewhat representative of an average actuary/actuary-in-training. The average financial success of Team Three is assuredly materially below that of Team One. Their intelligence is average (for actuaries) and not at the superior level that we assign to Team Two participants. Team Three's average age is likely lower than Team One's age and higher than the average age of Team Two. This may present some disadvantage with regard to general knowledge questions relative to Team One and with regard to pop culture questions relative to Team Two. The only profile we can compile is that of an average group of actuaries. Team Three is lacking the proven performance of Team One and the demonstrated intelligence of Team Two. Since they are an average group, we should expect average performance in Trivial Pursuit, making them the least likely of the three teams to win. Average must always lose out to proven individuals or teams with superior intelligence. Or is it possible that there are exceptions to this rule?

Since success in Trivial Pursuit is dependent on the collective result of all team members' wisdom, knowledge and experience, most readers of this article will readily conclude that in spite of their average profile, Team Three should be the odds on favorite to win at Trivial Pursuit. The strength of Team Three is in its diversity. This team most likely has members who are: young, seasoned and in between; female

and male; life, health and casualty professionals; varying degrees of intelligence; astute business professionals, strong technicians and those with developing skills; and a host of other characteristics that will most likely be lacking in the other, less diverse, teams. Trivial Pursuit does not require every team member to know the answer to every question. It simply requires that one team member know the answer to the question posed and that this individual express a conviction that their answer is correct. Because of their diversity, Team Three should be favored to win at Trivial Pursuit over the other two teams.

CROWDS MAKE BETTER DECISIONS

This Trivial Pursuit example amplifies the primary tenet of James Surowiecki's book that a diverse group of individuals, namely, a crowd, has greater wisdom and can make better decisions than a very accomplished individual expert or a very smart group of individuals who are similar to each another. This idea runs counter to many people's intuition and to the practices of many corporations. Company decision makers are often the heads of departments or heads of companies who often use the input of external experts or the recommendation of homogeneous groups of employees to make decisions. Surowiecki argues that decision making would be better left in the hands of a diverse group of employees.

Why is that which is obvious to Surowiecki not clear to the rest of us? Why do we place so much value on leaders as decision makers? Why do we value opinions of experts more than we value our own judgments? Perhaps the answer is as simple as the fact that we over-value individual intelligence and under-rate the wisdom of the masses.

Using examples and statistics, Surowiecki suggests that individual experts are considerably less effective at finding the best solutions or approaches than diverse groups. Experts have a great confidence in their own abilities and tend to overestimate the correctness of their solutions or decisions. Studies have shown that experts are as likely to disagree with the opinion of another expert as they are to agree. Although there may be some exceptions, chasing

a single expert opinion is generally a losing proposition. It would be far superior to gather the perspectives of several independent experts than to rely on the perspective of only one. Many companies choose to use the opinions of experts rather than accepting the average opinions of groups because they believe that averaging is equivalent to dumbing-down or compromising. Their prejudice and practice is towards finding the one right person to guide and direct them to the right decision. Companies who chose one expert based on past successes are, according to Nassim Taleb, fooled by randomness when they judge that the past success of any one individual is repeatable. More likely the prior success was a result of a string of good luck.

TO TAP THE WISDOM OF CROWDS, A GROUP MUST BE DIVERSE

If forming a group leads to better decisions, why would we expect Team Two, a group of highly intelligent individuals to lose to Team Three with their average profile? Because simply forming a group to make a decision is not enough. To tap the wisdom of crowds, a group must be diverse. Homogeneous groups are prone to "groupthink." Due to their commonality they quickly become cohesive. As they become cohesive they tend to become isolated from outside opinions and convinced of their own judgment. Because the group's beliefs are easily reinforced amongst themselves, these groups quickly sense the feeling of consensus and close out dissenting opinions. Members find it easier to change their opinion to conform to the group than to challenge the group with their independent and conflicting ideas. Surowiecki gives examples as to how this groupthink occurs in homogeneous groups and uses psychological experiments as evidence to support this.

In contrast to homogenous groups, diverse groups make it easier for individuals to say what they really think. In diverse settings, the varied opinion of each individual does influence the thinking of the entire group. Diversity allows individuals to maintain an independence of opinion that is critical for groups to function at an optimal level. The smaller the group, the more important it is to have diversity of thought, experience and perspective. Having intelligent and

informed members is important to any group but equally important is having members who may have less information but have different skills than the more knowledgeable members. Diverse groups not only generate a more diverse set of solutions than homogeneous groups; they are also better equipped to distinguish between good solutions and bad. A diverse group needs to operate without hierarchy if it is to produce the best decisions. Top down structure will more likely produce the answer provided or desired by the senior leader whereas a more decentralized group will maintain an independence of opinion. Decisions must not be dictated from a centralized source but rather must be made by the individuals in the group using their own local and specialized knowledge. Specialized knowledge must make its way into the decision-making process even though often the details of the specialized knowledge are difficult for the person who owns them to explain to others. I think actuaries understand all too well how their own specialized knowledge can be difficult to explain. Perhaps more important to the group decision making process than the actuary's explanation of their specialized knowledge is the actuary's description as to how their specialized knowledge should impact the final decision.

Another key element of good group decision making is arranging for people to offer their judgments simultaneously rather than sequentially. This eliminates a significant element of peer pressure that will stifle independence of thought. Every independent thought and perspective should be a part of the discussion of an issue. Based on his studies, Surowiecki goes so far as to suggest that encouraging people to make guesses and/or state their opinions based on their private information—even if their guesses or opinions are judged by others to be obviously incorrect—will actually make the group as a whole smarter.

THE USE OF THE DEVIL'S ADVOCATE

One of Surowiecki's ideas that I find especially appealing is the use of the devil's advocate to help groups make better decisions. Often teams formed to make a decision will start, either overtly or tacitly, with an expected conclusion. When this occurs, all information from that point forward is shaped to conform to the foregone conclusion. Generally the only information that will be discussed is information that everyone already knows. A devil's advocate can change the dynamics of this entirely by causing the group to pay attention to the available information without conforming it to the preconceived notion that supports the conclusion. Studies have shown that, especially in small groups, having even a single different opinion can make the entire group wiser.

Surowiecki discusses prediction markets in some detail in his book since these seem to epitomize how to tap into the wisdom of crowds. Prediction markets have an advantage over small groups in that their sheer numbers produce the diversity necessary for the wisdom of crowds to emerge. Prediction markets allow large numbers of independent investors to make bets about the likelihood of various events occurring. Sports-betting is already well established and is fairly efficient at predicting outcomes of many sporting events through the collective intelligence of those betting. Following this model, prediction markets have already been established to tap into the wisdom of crowds on a variety of topics including predicting outcomes of elections, which movies will be successful and the expected period of success of political leaders. The potential of prediction markets has barely been tapped and represents a great tool for actuaries for forecasting and gathering information, both currently and in the future.

Surowiecki's book is full of interesting examples of how groups function. Not all groups produce superior results and studying what doesn't work with groups is as important as studying what does work. Surowiecki gives numerous examples how groups make bad decisions. He describes in fascinating detail various groups gone amuck including:

• How the bowling investment bubble formed and popped in the 1950s and 60s.
• How traffic jams form.
• How poor group process led to bad decisions in the Bay of Pigs.

• How bad decision making and group dynamics led to the Columbia Shuttle tragedy.

Other examples Surowiecki spotlights make informative case studies of group dynamics including:

• How within hours Wall Street pointed to the most likely contributor to the Challenger explosion six months before the Presidential Commission did so.
• Why people tip when there is no personal benefit to doing so.
• Why capitalism requires impersonal trust.
• How Google uses the wisdom of crowds to produce search results in fractions of seconds.
• How a naval officer assembled a team with a varying background and offered prizes of Chivas Regas to accurately pinpoint the location of the missing U.S. submarine Scorpion in 1968 when the only available information was the Navy reports of its last known location.

SPREADING THE WORD ABOUT THE WISDOM OF CROWDS

One interesting result of my reading T*he Wisdom of Crowds* is that every day since finishing the book I have been mentioning, discussing and applying what I have learned with my coworkers. The potential opportunities to apply the learnings from this book are present in life every day. From a practical perspective, this demonstrates that the ideas of this book have a great utility and can impact the way we interact with others in our daily living. Because of my personal experience in discussing and using the insights from my reading, I highly recommend this book.

I'd like to extend the Trivial Pursuit example given to start this article. It might be expected that Team Four, a team of 10 randomly chosen adults from North America would outperform Team Three, North American Actuaries, in Trivial Pursuit even though—and forgive my arrogance—the average intelligence of the North American Cross-Section Team would likely be less than the average intelligence of the North American Actuaries Team. The greater diversity of the North American Cross-Section Team would most certainly give them an advantage relative to the somewhat insular professional homogeneity of the North American Actuaries Team in a game of Trivial Pursuit. However, winning at trivial pursuits should be considered small stakes. The big stakes for actuaries involve the decisions we make in our businesses. Based on the information presented in *The Wisdom of Crowds*, you should ask yourself the following questions:

• When you are involved in a decision-making meeting, are you meeting only with other actuaries or are you meeting with a wide range of interested parties?
• Are all voices and perspectives being expressed and processed?
• Is anyone expressing dissent?
• Is one person suppressing or shaping the opinions of others?
• Are you sharing your perspective, your specialized knowledge?

By understanding the wisdom of crowds and the importance of diversity in the decision-making process, you can not only extend your influence by making your opinions heard but you can also improve the decision-making process for your entire organization by making certain that the opinions of a diverse group of thinkers are part of a group-based decision-making process. I hope you'll take the time to read *The Wisdom of Crowds* to better understand what it takes to allow groups to make better decisions. ▼

Scott McInturff

Scott McInturff, FSA, MAAA, is vice president—Individual Insurance Product Management of the U.S. Operations—Sun Life Financial in Wellesley Hills, Mass. He can be reached at *scott.mcinturff@sunlife.com.*

Complexity:
A Guided Tour
by Melanie Mitchell, Ph.D.

Reviewed by Dave Snell
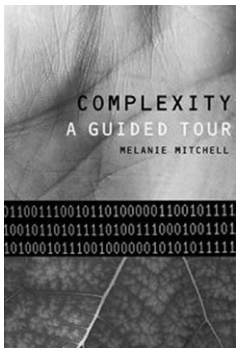
One of my favorite books from the early 1980s was a huge tome titled *Gödel, Escher, Bach: An Eternal Golden Braid,* by Douglas Hofstadter, a pioneer in Artificial Intelligence (AI). Hofstadter described GEB (the initials became a popular abbreviation for his book) as "a metaphorical fugue on minds and machines in the spirit of Lewis Carroll." At the time I was just getting interested in AI and I found GEB fascinating. Apparently, I was not alone. Melanie Mitchell, then a high school mathematics teacher in New York, found it "one of those life-changing events that one can never anticipate."

She wrote to Hofstadter indicating she wanted to study under him as a graduate student. Receiving no reply, she later approached him in person when he gave a lecture at MIT. He handed her off to a graduate student. She was "disappointed, but not deterred" and after several more follow-up calls to him, she managed, through her persistence, to convince him of her passion for AI—a topic that eventually was absorbed into complexity science. Eventually, she moved to Michigan and earned her Ph.D. under Hofstadter and John Holland, another complexity science pioneer. I mention this history to try to convey the contagious enthusiasm for complexity science that Dr. Mitchell exudes in her book. She seems to prefer the term complexity sciences, since this is such a cross disciplinary subject; but in this review I'll use the more common term, complexity science.

Mitchell starts with an acknowledgement to the Santa Fe Institute (SFI) where she directed an SFI Complex Systems Summer School. The SFI seems to be the current epicenter for complexity science research, and this book is an expansion of the author's series of SFI lectures on "The Past and Future of the Sciences of Complexity," with updated material reflecting new perspectives from 2008 and 2009.

Previous knowledge of complexity science is unnecessary, as the first chapter starts out with a series of examples to describe what is meant by complexity. This was useful since the topic seems to evoke many different definitions from scientists and practitioners. An actuary often likes to start with some definition of the topic under study; but a rigorous and widely accepted definition of complexity science just does not exist yet. On the other hand, we are eagerly embracing the study of enterprise risk management (ERM) to the point of promoting the Chartered Enterprise Risk Analyst (CERA) designation as an international credential of expertise in risk management—even though we may differ considerably in our opinions about what constitutes risk. In a similar vein, Mitchell's examples make it clear what falls into the realm of complexity. The examples run the gamut from insect colonies to the human brain; and from immune systems to economies and the World Wide Web. In some respects, ERM seems like an application of complexity science; and quoting A.S. Eddington, the astronomer who first demonstrated that Einstein's Theory of Relativity worked in the real world, "We need scarcely add that the contemplation in natural science of a wider domain than the actual leads to a far better understanding of the actual." I submit that a study of the wider domain of complexity science can help us better understand risk management. In fact, lest the actuary reading this assume that the book mentions only theory and some science applications, the author peppers her theory with references to practical financial applications in several sections. She explains early on that:

> Economies are complex systems in which the "simple, microscopic" components consist of people (or companies) buying and selling goods, and the collective behavior is the complex, hard-to-predict behavior of markets as a whole, such as changes in the price of housing in different areas of the country or fluctuations in stock prices.

and later in the book she gives specific examples:

GAs [Genetic Algorithms] have been used by several financial organizations for various tasks: detecting fraudulent trades (London Stock Exchange), analysis of credit card data (Capital One), and forecasting financial markets and portfolio optimization (First Quadrant).

Her extensive notes section refers the reader to details about each of these specific applications.

In *Complexity: A Guided Tour*, we are given a short history lesson on the roots of dynamical systems theory, chaos and prediction. Again, the examples help guide the reader through an inductive learning process. Deterministic chaos, for example, is introduced via the famous logistic map that results from varying values of R in the seemingly simple equation $x_{t+1}=R \cdot x_t \cdot (1-x_t)$ where $0 \leq x_t \leq 1$. Along the way, we hone in on Feigenbaum's constant, a universal constant for functions approaching chaos via period doubling, and the fact that it applies outside the realm of pure mathematics and shows up in electronic circuits, lasers and chemical reactions.

Now, we are ready to approach the concepts of information, energy, work and entropy. This is explained through stories about the development of the Second Law of Thermodynamics, Maxwell's Demon and Shannon's Information Theory. Moving along to computation, Mitchell guides us through topics such as "What is Computation and What Can Be Computed?" She describes Hilbert's Problems and Godel's Theorem, which proved that not all mathematical questions are computable. Then she covers Turing machines, where the goal is to mimic human behavior so well as to fool a human, and this leads into a chapter on evolution. Her primer on evolution summarizes pre-Darwin, Darwin, Mendel and the Modern Synthesis, and leads quite naturally into the next chapter, on Genetics. Skipping quickly through an admittedly simplified treatment of DNA and RNA, Mitchell leads us into

Dave Snell

Dave Snell, ASA, MAAA, is technology evangelist with RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *dsnell@rgare.com*.

the geometry of fractals, and the underlying power laws that describe them when normal measurement techniques fail us.

Now with approximately 100 pages of history and basic tutorials behind us, we can begin the next parts of the book, which deal with topics such as life and evolution in computers, cellular automata, information processing in living systems, genetic algorithms, ant colony optimizations, and the mystery of scaling. Clearly oriented towards AI, the author devotes a major chapter to applying network science to real-world networks—such as the brain.

Each topic is approached in a logical, understandable manner. In addition though, as a reader I felt the excitement of the discovery process as I read about Von Neumann's self-reproducing automation, the "New Kind of Science" from Wolfram, and the gradual increase in intelligence of Robby, the soda-can-collecting robot, like the movie robot *WALL·E,* which eventually outperformed the author in developing its own clean-up strategy.

A chapter is devoted to an overview of the author's development of "copycat"—a program she wrote for her doctoral dissertation that makes analogies in the letter-string world by using reasoning believed similar to that used by humans as we make analogies to understand our world.

The examples often caused me to stop and write a quick spreadsheet or program to further explore the particular subject.

This is one of the first books I read on complexity science; and admittedly many of the ones I read afterwards were more narrowly focused; and some went into more detail, or provided even more memorable examples on particular complexity science topics. However, this book gave me a base level understanding of a lot of topics that previously were just fancy sounding phrases. More than that, it nurtured my initial interest in complexity science and left me with a voracious appetite for more! The subtitle is appropriate. This is truly a *guided* tour for complexity. Dr. Mitchell is an excellent guide; and I recommend her for your visit to the amazing world of complexity science. ▼

Forecasting Judgment: The Netflix Prize and Collaborative Filtering

By Mike Lindstrom

One way to predict the numerical outcome of an event is to ask a panel of experts, and then take the average of their answers. Delphi studies, statistical surveys and other judgmental forecasting techniques are commonly used when attempting to glean the general from a variety of individual opinions. Some forecasters however, move in the opposite direction. Can the aggregate results from a group be used to forecast something as subjective and judgmental as one single person's opinion? Forecasting of this type commonly uses a set of techniques known as collaborative filtering, and one company has paid $1,000,000 to do it better.[1]

THE NETFLIX PRIZE CHALLENGE

Beginning in 2006, Netflix sponsored a contest to "predict whether someone will enjoy a movie based on how much they liked or disliked other movies."[2] At that time, the company currently had a predictive model, in order to recommend movies to individual customers. The goal of the contest was to improve on their current model by 10 percent.

More specifically, the challenge was to design a model that can predict the rating a particular user would give to a movie. Netflix users are allowed to give movies a rating of one to five stars. Teams of contestants designed models that predicted the ratings for particular users and movie titles. These modeled ratings were then compared to the actual ratings the user had given to those movies. How well a model performed was gauged by comparing the root mean squared error of model results with the root mean squared error of the Netflix model on the same data set.

In addition to the grand prize, Netflix also offered an annual $50,000 progress prize, which was awarded in 2007 and 2008 to the leading team. By October of 2009, the contest had been won when a team beat the current prediction algorithm by 10.06 percent.[3]

USING A MODEL TO JUDGE: COLLABORATIVE FILTERING

How can a model predict which movies I will like? I really enjoyed the movie "Fargo," as well as "The Big Lebowski," two movies directed by the Coen Brothers. When they released another movie called "No Country for Old Men," I was pretty sure that I would like that movie as well. This is called the nearest-neighbor approach and is a type of forecasting called collaborative filtering. In the nearest-neighbor approach, a piece of the predicted rating is a weighted average of similar, or neighboring items, where the weights are the item similarities.[4]

A more abstract type of collaborative filtering is called the latent-factor approach. This approach is used to discover patterns in the data set that may not be apparent or easily explainable. The predicted review is then based on a set of factors that correspond to the particular movie and user.[5]

DIFFICULTIES IN MODELING JUDGMENT

Obviously at this point human judgment and opinion cannot be completely captured in a mathematical model. For example, an average answer might be way off if opinions are at the extremes. Consider a problem discovered early on in the Netflix contest called the "Napoleon Dynamite" effect,[6] which arises from the movie of the same name. This is an example of a movie that people tend to either love or hate, so actual ratings are either a "1" or a "5." Any model that relies on an average prediction for that type of movie will show a relatively large error each time.

There are other undeniably human attributes that lead to inaccurate predictions, some of which can be modeled.[7] Our opinions are significantly influenced by things that are within close proximity, a bias in judgment called anchoring. For example, I might watch several really bad movies and then give a good movie a lower rating than it deserves. Also, temporal effects are evident in ratings. Types of movies can go in and out of style. Other factors can affect reviews like the day of the week a movie was watched, or the number of movies that had been watched recently.

Also, no two people use the same absolute scale so a model could take into account the fact that some people are harsher critics than others. Some users may reserve five stars for only their absolute favorite movies while other people may concentrate ratings on the high end of the scale.

OTHER FORECASTING COMPETITIONS

Besides the Netflix Prize, other organizations have sponsored their own forecasting competitions and challenges. Most notably these include the M Competition and the time-series prediction competition sponsored by the Santa Fe Institute.

In this edition, you can see information on a forecasting application contest sponsored by our section. The award is not as significant as the Netflix Prize, but we urge you to enter and share your thoughts and expertise. Later this year the Forecasting and Futurisim Section will sponsor a larger forecasting competition for actuaries. Look for an announcement of the competition at the SOA Annual Meeting.

Mike Lindstrom

Mike Lindstrom, ASA, MAAA, is associate actuary at Federal Life Insurance Company (Mutual). He can be reached at *mlindstrom@federallife.com*.

REFERENCES

[1] Netflix, Inc., "Netflix Prize: $1 Million Paid, Another Million Now Up For Grabs" *http://blog.netflix.com/2009_09_01_archive.html*

[2] Netflix, Inc., "Netflix Prize Rules" *http://www.netflixprize.com//rules*

[3] Netflix, Inc., *http://www.netflixprize.com//leaderboard*

[4] Bell, R., and Koren, Y., "Lessons from the Netflix Prize Challenge," SIGKDD Explorations, Volume 9, Issue 2, 2007

[5] Bell, R., Bennett, J., Koren, Y., and Volinsky, C., "The Million Dollar Programming Prize," *IEEE Spectrum, May 2009.*

[6] Thompson, C., "If You Liked This You're Sure to Love That," *New York Times Magazine,* 21 November 2008

[7] Whimsley (blog), found at *http://whimsley.typepad.com/whimsley/2009/10/netflix-prize-was-the-napoleon-dynamite-problem-solved.html* ▼

A New End, A New Beginning
Preparing for Life as We Don't Know It

By John L. Petersen

Predicting the future is a fool's errand. It is fraught with so much complexity and uncertainty that the best one can do with integrity is to array potential alternatives—scenarios—across the horizon, and then try to think about what might be done if one of those alternative worlds materializes.

Scenario planning has certainly been an effective discipline, helping many organizations to imagine potentialities that probably otherwise wouldn't have shown up in their field of view. But as I facilitate organizations going through these exercises, the little, nagging voice in the back of my head is not asking, "What is the array of possible futures?" Rather, it is always wondering, "What is the future really going to be?" We want concreteness. We want predictions.

I think that no one knows for sure what the future will bring, but after some time of being in the "future business," one begins to be able to discriminate between what is substantive and structural and what is largely speculative. For me, at least, some things have an intuitive sense of being real and important, and the rest of the possibilities lack just enough gravitas that I know they're only "ideas." That intuitive sense is supported when it becomes possible to triangulate from a number of independent sources that all point to the same conclusion.

People always ask me after my talks, "With all of these converging trends, what is 2012 really going to look like?" It happened again in a recent radio interview. Mostly I hedge and dance a bit and say that I don't know for sure, but I believe there will be a new world, and a new human will come out of all of the current turmoil. The notion of cooperation will shape the way people see themselves and the rest of the world, and there will be new institutions and functions, etc. Pretty general stuff.

But, over a year ago, the notion that all of this big change could spell a substantial reconfiguration of the familiar country that I have lived in all of my life began to gel in a way that moved that notion beyond being just a possibility—a wild card—into the realm of plausibility. I now have come to believe that such a transformation is likely and will happen—soon.

Ideas like this are so big and disruptive that it is really quite hard to get to the place where we take them seriously. For most of us, our lives are evolutionary—punctuated, perhaps with trauma now and then, but mostly populated by events that are familiar, even if they don't always make personal sense. The concept that everything might change is so foreign to any experience that most of us have ever had that, even if we say the words and talk about the possibility, we really don't internalize what this might mean.

Certain other thinkers jumped to the natural conclusion quite some time ago. Dmitry Orlov, for example, first started to build a theory of superpower collapse that included the United States in 1995. Only in the last few years has he been talking publicly about his ideas and the ultimate direction of U.S. trends.

James Howard Kunstler, a wonderfully entertaining and provocative writer, was very clear about the systemic and structural nature of the larger problem in his 2006 book, *The Long Emergency*. He clearly sees the demise of America coming this way. His always interesting blog is a weekly assessment of where we're going wrong.

My colleague David Martin outlined the financial dominoes that were going to fall in a talk at The Arlington Institute in July 2006. Implicit in his treatise is the collapse of the U.S. and global financial systems, but again, it's one thing to hear such views and quite another to really believe them.

After I listened to such people and pondered what they said, I began telling my friends that I thought we were seeing the beginning of the end of the United States as we've known it. I didn't think they really believed it, at least initially, but recently we have seen Singapore, for instance, reportedly making major leadership changes in its government investment company to reposition the nation away from the United States and the U.S. dollar.

INDICATORS OF BIG CHANGE AHEAD
There are numerous indicators that suggest the big change is coming:

• **Multiple trends are converging.** Huge, extraordinary, global trends, any number of which would be enough to derail our present way of life, are converging to precipitate a historic big transition event. A partial list would include:

- The global financial system is collapsing. During the next few months, it appears that wave after wave of blows will strike the system, raising the very real possibility that it will experience large-scale failure sometime before the end of the year.
- We have reached the beginning of the end of petroleum. Global production has been flat for the last three years. Senior oil company executives are now saying that they will not be able to pump more. Supply will likely begin to decrease significantly after we move across the peak. Prices will increase again if the demand holds up. This is important because our present way of life is built upon petroleum.
- The global climate system is changing—some say it is getting much warmer; others now suggest a mini ice age within the next decade. In any case, increased irregularities in local climates will probably result, with attendant problems in agriculture, natural disasters and economies.
- The cost of food is increasing rapidly as a result of global shortages not seen in 40 or 50 years. This could be exacerbated by increasing energy costs and climate changes. Lester R. Brown of the Earth Policy Institute believes that food shortages may bring down civilizations.

"THE END IS NEAR" HAS ALWAYS BEEN DOOMSAYERS' FAVORITE SLOGAN, BUT IS IT NOW FINALLY TRUE? THE TRENDS SUGGEST THE END OF AN ERA MAY INDEED BE NEAR, AS GROWING COMPLEXITY AND PROLIFERATING CRISES THREATEN TO OBLITERATE "LIFE AS WE KNOW IT." THE TIME IS NOW TO PREPARE FOR THE LIFE WE DON'T YET KNOW.

- The effects of larger solar eruptions hitting the earth through tears in the magnetosphere surrounding our planet will likely disrupt global communications, weather, perhaps satellites, and even organic life over the next three to four years.

• **Problems are much larger than government.** Peak oil, climate change, and the financial meltdown all have the potential to significantly overwhelm the capabilities of government to respond to them. If bureaucracies can't deal with the aftermath of a natural disaster like Katrina, something 10 or more times that damaging would leave most people fending for themselves. If these extraordinary, disruptive events end up being concurrent, then the whole system will be at risk.

• **The problems are structural.** They're systemic. Some of these issues, especially the financial, oil, and food problems, are also a product of how we live, our priorities, and our paradigms. We are creating the problems because of our values and principles. Without extraordinary, fundamental changes in the way we see ourselves and the world, we will keep getting what we are getting.

• **Leaders think the old system can be "rebooted."** Almost everyone in leadership positions in the Obama administration and in other countries wants to make the old system well again. Jim Kunstler has said it well:

Among the questions that disturb the sleep of many casual observers is how come Mr. O doesn't get that the conventional process of economic growth—based, as it was, on industrial expansion via revolving credit in a cheap-energy-resource era—is over, and why does he keep invoking it at the podium? Dear Mr. President, you are presiding over an epochal contraction, not a pause in the growth epic. Your assignment is to manage that contraction in a way that does not lead to world war, civil disorder or both. Among other things, contraction means that all the activities of everyday life need to be downscaled including standards of living, ranges of commerce, and levels of governance.

"Consumerism" is dead. Revolving credit is dead—at least at the scale that became normal the last 30 years. The wealth of several future generations has already been spent and there is no equity left there to refinance.

The above indicators of change suggest the reasons behind the following.

• **We're not dealing with the structural issues.** All of the biggest efforts are attempts to re-inflate the financial bubble and to keep the mortally wounded institutions alive. The knee-jerk reactions come from the same people who helped to design and feed the present system. These people are also deluded—they think (or act like) they know what they are doing. They don't realize that …

• **The situation is so complex that no one really understands it.** The Global Business Network's Peter Schwartz, reporting on a conversation with the *Financial Times's* Martin Wolf, said that Wolf's key point was that the nature and scale of the credit crisis is so novel that it's not clear we know what we're doing when we try to stop it. He is deeply worried. Steve Roach of Morgan Stanley said at the World Economic Forum annual meeting at Davos that he agreed with Wolf: We are in uncharted waters. Nassim Nicholas Taleb, author of *The Black Swan: Impact of the Highly Improbable* (Random House, 2007), says the financial system is so complex that it is impossible for anyone to understand it, and because of that complexity it is inevitable that it will exhibit significant, unanticipated behaviors (his Black Swans) that career across the planet.

• **The issues are global.** Economies are contracting around the world, with a huge rise in unemployment. Japan's exports are falling, and factories are closing in China, which means that products aren't being shipped.

• **The system is fundamentally out of balance.** Common sense is largely absent from many big, sweeping U.S. government edicts. The Transportation Security

HUGE, EXTRAORDINARY, GLOBAL TRENDS … ARE CONVERGING TO PRECIPITATE A HISTORIC BIG TRANSITION EVENT.

Administration, for example, wants to make pilots produce background checks on members of their family (and their business associates) in order to legally give them rides in noncommercial, private airplanes. The Agriculture Department wants all small farmers to put GPS/RFID tags on all of their animals so that chickens, cows, horses and goats can be tracked, on a day-to-day basis, by the government. And most of the U.S. federal budget goes to the military and military-related agencies. This kind of growth, of course, is what brought down the Soviet Union.

WHY GOVERNMENT FAILS TO RESPOND TO CHALLENGE

If the natural solutions to these massive issues include innovation, foresight, adaptability, sustainability and resilience, it is unlikely that a thinking American could be found who would suggest that the source for these capabilities would be our government. Those who are in charge have no new ideas about how this all should work. They're also slow, and this situation needs fast, agile responses. There is an additional problem. Even if it did have good ideas, the government wouldn't be able to effectively implement them because:

• **It suffers from too much inertia, and too many lawyers and lobbyists.** There is a huge, well-funded effort in place to maintain the status quo or to shift the future to benefit one group at the expense of others. It would be impossible within the present system to initiate dramatic change when the threat was still on the horizon. Every group or organization that might be negatively affected would fight in Congress and the courts to keep themselves alive, regard-

less of what was at stake for the larger community. Only when the crisis was about to crash down on everyone—when adequate time and resources for effective response were nonexistent—might everyone pull together for the common good.

• **Potential solutions take too long to implement.** These issues are so gigantic that confronting and redirecting them takes a long time. One study, for example, suggested that a national crash program to find alternatives for oil would need to have been started 20 years before the peak in order for there not to be significant disruption of the underlying systems. We do not operate with either that foresight or that resolve.

• **Supply chains are long and thin.** Globalism and just-in-time production have produced supply chains in most areas of commerce that are very long—often to the other side of the earth—and very fragile. There are many places between there and here where something can go wrong. If and when that happens, necessities will not be available. In those situations, people resort to unconventional and/or antisocial behavior.

• **Six hundred trillion dollars in derivatives is a house of cards.** Looming over the whole financial situation is an almost unfathomable quantity of financial instruments—derivatives—which are essentially casino bets with no underlying value supporting the transaction. Warren Buffett calls them "financial weapons of mass destruction," and they could bring the whole system down. Derivatives only work if there is confidence in the system—you believe the casino will really pay your winnings. If other things in the environment erode that confidence, there is the real possibility that things will rapidly fall to pieces.

John L. Petersen

John L. Petersen is president of The Arlington Institute, a nonprofit, future-oriented think tank. Among his books are WFS best-sellers, *The Road to 2015* (1994), *Out of the Blue* (1997) and *A Vision for 2012* (2008). He can be reached at *johnp@ arlingtoninstitute.org.*

• **Cooperation is unlikely; protectionism will prevail.** Instead of countries cooperating with each other to deal with these big transnational problems, we're seeing a pulling back to protect each country's perceived short-term interests, regardless of what the implications might be in the longer term. At the same time, we're all connected to each other in very complicated ways, so if any substantial pieces of the system don't work, it will affect all of the other ones.

• **History says it's time.** Perhaps what is most compelling to me is that history strongly suggests that the time is right for an upset—they always happen about now in the historical cycles. Big punctuations in the equilibrium of evolution have produced extraordinary, fundamental reorganizations to life on this planet on a regular, accelerating basis from the beginning of time as we know it. We make progress as a species when we are forced in one way or another to evolve into seeing ourselves and the world in new ways. Necessity is the mother of invention—or should be!

So, it doesn't look to me like we're going to be able to do what might be needed to maintain the present system. It is likely that we're at one of those extraordinary moments in history when each of us gets the opportunity to play an important role in not only transitioning to a new world, but also designing it.

WHAT TO DO IN THE FACE OF UNPRECEDENTED CHANGE

Two specific actions come to mind that should help individuals and institutions prepare for this change.

1. Plan for the transition. Start to think now about how you're going to provide for yourself and those who are important to you in a time when many things don't work the way they always have in the past. There are many websites and books on this subject, but the key concept is cooperation. You can't do this alone. Start to work together with like-minded individuals to sustain yourself, regardless of whether your concerns are food, water, shelter, transportation or finances.

2. Start thinking about the new world. Now is the time to begin contemplating the design of the new world. Governments should be doing this. Companies should start skunk works. Big international organizations should put it on their agendas.

Here's the catch. This might not happen. The "system" might not collapse. Personally, I think that if there is any one person that has the potential to at least soften this transition it is Barack Obama. As I've suggested, he will have his hands full just trying to get the underlying people and institutions to think differently and act fast enough, but if anyone has the chance to pull it off, he would be the one. Already, he's getting government to move faster and in more substantive ways than any of his predecessors. It may be, by the way, that he will be the best person to wind down the old system and develop a new one. It's all of the other folks running the government that I'd be concerned about—the ones who continue to see the world as it used to be.

There are any number of reasons why this scenario might not manifest itself, not least of which is that there will be many thousands, if not millions, of people who will be working very hard to assure that the system doesn't come apart (but then, they may be doing the wrong things).

It seems to me, therefore, that flexibility and permeability (allowing new ideas to get through) are of critical importance here. Remember the first law of Discordianism: "Convictions cause convicts." Whatever you believe imprisons you.

So, stay loose. The winners need to transcend, not try to work their way through all of this. Concentrate on building the new world. Don't get emotionally involved in the daily reports of the current global erosion.

ADDITIONAL RESOURCES
• *America's Defense Meltdown: Pentagon Reform for President Obama and the New Congress*, edited by Winslow Wheeler (Stanford Security Studies, 2009). In sharp contrast to the political apparatchiks protesting that more money is needed to reverse the shrinking, aging, and declining readiness of the Army, Navy, and Air Force, few seem to understand that budget increases are a primary cause of the problems, a symptom clearly described in this new book.
• "Asymmetric Collateral Damage: Basel II, the Mortgage House of Cards, and the Coming Economic Crisis," a talk by David Martin, CEO of M·CAM, for The Arlington Institute (July 12, 2006), *www.arlingtoninstitute.org/dr-david-e-martin.*
• Crash Course, a "concise video seminar on how our economy, energy systems, and environment interact, and how they will impact the future," at *www.chrismartenson.com.*
• *Reinventing Collapse* by Dmitry Orlov (New Society, 2008).
• *A Vision for 2012: Planning for Extraordinary Change* by John L. Petersen (Fulcrum Publishing, 2008).
• "Why Obama's 'Rescue' Misses the Mark and the Coming Financial Collapse Just Got Worse" by David Martin of M·CAM (February 15, 2009), *http://invertedalchemy.blogspot.com/.*

(Reprinted from the September October 2009 issue of *The Futurist* [the WFS magazine]. Used with permission from the World Future Society, 7910 Woodmont Avenue, Suite 450, Bethesda, Maryland 20814 USA. Telephone: 301.656.8274; *www.wfs.org*) ▼

Living with Actuarial Black Swans—a Discussion with Nassim Nicholas Taleb

By Ben Wadsley

I may be killed crossing the street; should I not cross? Taleb's answer is, "Do not cross the street blindfolded!"

At the 2009 SOA Annual Meeting in Boston, Nassim Nicholas Taleb shared some thoughts on Black Swans, operating in the fourth quadrant, and living in the world of Extremistan versus Mediocristan during his keynote speech, all of which can be found in his published work. The follow-up session, summarized here, was our chance to ask the real question: What do we actuaries do about these phenomena?

The panel discussion was jointly sponsored by the Financial Reporting, Investment, and Forecasting and Futurism Sections and featured presenters Steve Conwill, FSA, MAAA, Max Rudolph, FSA, CERA, MAAA, and John Tiller, FSA, CERA, MAAA, respectively. The session was moderated by Alan Mills, FSA, ND. Look for other great sessions from these three sections at the 2010 SOA Annual Meeting in New York. The panelists asked several questions from their respective actuarial fields' point of view, and there were three recurring topics from Mr. Taleb's responses that are summarized below: Disclosure, Stress Testing, and Globalization.

DISCLOSURE

In 2003, Taleb was quoted in *The New York Times* saying that Fannie Mae will go bust. A disgruntled former employee of Fannie Mae had provided a copy of an undisclosed risk management report to a *Times* journalist. The journalist took it to Taleb, who was a professor at NYU, for interpretation, leading to Taleb's prediction. Fannie Mae had this risk report, but without disclosure the investors of the company and the media were not aware of the issues, thus they were not required to act on the results.

Ben Wadsley

Ben Wadsley, FSA, MAAA, is at AEGON USA, Inc. in Cedar Rapids, IA. He can be reached at *bwadsley@aegonusa.com*.

Are banks today marking-to-market their loans? No. In derivative trading where mark-to-market is used, every morning you start fresh. This avoids anchoring bias, which is the human tendency to rely too heavily on one piece of information when making decisions. In contrast, banks keep their loans at full value, even when payments are past due—the banks are anchored to the book value of the loan. The solution to this problem is to use mark-to-market reporting and have volatility instead of having nothing and then experiencing a Black Swan event.

Taleb went as far as to say the banks should stop hounding customers with late payment notices, realize the loss, and renegotiate the loans. This viewpoint may be a bit extreme, but the point he was stressing was that forcing mark-to-market should be extended past where it is being used today. A natural response to forcing banks to mark-to-market is the concern that it would cause unnecessary panic to shareholders. In the age of the Internet and the 24-hour news cycle, rumors spread quickly. Instead of mitigating rumors, be

robust to the rumors. This means that we should have more transparency through risk reporting and valuation so rumors won't have a chance to impact our business. With full disclosure, we can create robustness that will mitigate fragilities inherent in our complex economic systems. Disclosure won't directly make banks robust, but it will force banks to deal with issues immediately and directly which will make them more robust. According to Taleb, mark-to-market in hedge funds may "… make them seem more volatile, because everything is volatile. It's sort of like someone seems sicker because you take their temperature."

STRESS TESTING

To get real risk management value out of stress testing, extreme scenarios must be used. Smaller levels of shocks may not give us the true level of risk in an entity—which doesn't get us to the goal of stress testing—to measure a company's relative fragility and robustness.

Let us consider two portfolios. The first contains 95 percent Treasuries and 5 percent high risk securities. The second portfolio contains senior notes of a fully capitalized synthetic CDO. If only a moderate shock is used—as in early 2009 when banks were forced to stress test their business with an unemployment rate of 10.3 percent, which Taleb claimed was only a blink away from happening—the first portfolio may lose most of the value of the high risk securities, while the value of the second portfolio may hardly move. This would show that the first portfolio is riskier, when it is clear that it is not. A good stress test would have shown the robustness of the first portfolio (complete loss of the risky securities, but the Treasuries remained untouched) while the second portfolio would show much greater losses, revealing the fragility of the assets, thus filling the tail of the loss distribution.

It may be conceptually difficult to choose a stress test level greater than a historic high, but no largest historical high has a predecessor—World War I and 9/11 were many times larger events than any previous event.

GLOBALIZATION

The world is moving towards the interdependence and connectivity of a globalized world. In just the last year we have been presented with new challenges and have had to deal with terms like "too big to fail." One large loss that Taleb attributed to the impact of globalization was the 2008 Societe Generale loss of $7 billion caused by a rogue trader. The single rogue trader was hiding a $75 billion risk position, and the loss occurred when these contracts were unwound at fire-sale prices. Rogue traders cannot be prevented, but we shouldn't have a bank big enough to be able to take on that big of a risk position because of the nonlinearity in the risk taking ability of banks relative to their size. Taleb asked us to consider a thought experiment—if instead of one large bank with one rogue trader and a $75 billion

> SMALLER LEVELS OF SHOCKS MAY NOT GIVE US THE TRUE LEVEL OF RISK IN AN ENTITY—WHICH DOESN'T GET US TO THE GOAL OF STRESS TESTING—TO MEASURE A COMPANY'S RELATIVE FRAGILITY AND ROBUSTNESS.

risk position, there were 10 smaller banks with 10 rogue traders and a $7.5 billion risk position each. While a sales order of $75 billion caused a market movement of several percent, an order of $7.5 billion would only take two phone calls and would hardly cost anything on a liquid day. Taleb's prediction was that the loss may have been only $1 billion if the losses were spread across these 10 smaller banks.

Does Mother Nature already know this phenomenon? Perhaps the reason that we don't have a land animal bigger than an elephant is the inability of a large animal to survive Black Swans. In the case of an extreme drought, large animals requiring a large amount of water per day will die, while some—but not all—smaller animals requiring

less water will survive. This pronounced contagion effect applies to viruses as well. The unnecessary global travel that we do would increase the impact of a flu pandemic. When asked about the possibility of a flu pandemic, Taleb predicted that there is a "95% probability it will happen in my lifetime," and since there is a high probability that it will kill him, he is going to hurry and finish his next book, *Tinkering*.

Today's use of the Internet also redefines the "run on the bank" scenario. There was a value to the bank of having people stand in line to withdraw their money. In a modern-day equivalent scenario, everyone would have their transactions completed on their Blackberrys in minutes.

The combination of increased flu pandemic contagion effect and our financial interdependence makes it difficult to protect our businesses. Without dividing into smaller financial entities or subcontracting our risk-taking and reducing our global travel, Taleb points out that we are forced into buying "… insurance on the Titanic from a guy on the Titanic."

CONCLUSION
"There are so many errors we can no longer predict, what you can predict is the effect of the error on you!" — Nassim Nicholas Taleb ▼

SOA 2010

Annual Meeting & Exhibit

Oct. 17-20, 2010
Hilton New York
New York, NY

Join us at the **SOA 2010 Annual Meeting & Exhibit: Fresh ideas. Innovative seminars.** Top-notch speakers. Plus—oodles of networking opportunities. We're heading to New York—a city on the move. Vibrant. Electric. And we plan to infuse that energy into this year's meeting.

Check out what we have in store:

- General Session Keynote Speaker Ted Koppel—yes, we said Ted Koppel—the guy who has given us years of insight into the biggest news stories of our time and who is currently offering his take on current events as an NPR news analyst.

- Presidential Luncheon Keynote Speaker Paul Embrechts, the risk management guru who, when he's not working as professor of mathematics at the ETH Zurich, keeps busy working on international advisory groups, consulting to major financial institutions on quantitative risk management issues and authoring books like "Quantitative Risk Management: Concepts, Techniques and Tools."

- more than 100 sessions offering you professional development opportunities and greater interaction, participation and takeaways.

- cutting-edge research presentations and discussions.

- networking opportunities to make those important connections.

- exhibitors giving you a look into their latest and greatest offerings.

- exclusive sponsorship opportunities, providing your company a chance to connect with a valuable target audience.

Be there—in the city that doesn't sleep. Learn more at www.soaannualmeeting.org.

Forecasting
& Futurism
N E W S L E T T E R