Borrower: PSC

Lending String: *PZI,CSC,TXJ,UMR,IGA

Patron: Walter Rees Stromquist (SC)

Journal Title: Journal of petroleum technology ; official monthly publication of the Petroleum Branch, American Institute of Mining and Metallurgical Engineers.

Volume: 28 **Issue:** AUG
Month/Year: 1976 **Pages:** 843-850

Article Author:

Article Title: CAPEN,E 'DIFFICULTY OF ASSESSING UNCERTAINTY'

Imprint: [Dallas, Tex. ; AIME], 1949 1978

ILL Number: 41424559

Call #:

Location:

MF NOS

ARIEL
Charge
Maxcost: $50IFM

Shipping Address:
ILL
Swarthmore College Library
500 College Ave.
Swarthmore, PA 19081

Fax: 610-690-5782
Ariel: 130.58.88.95

The Difficulty of Assessing Uncertainty

E. C. Capen, SPE-AIME, Atlantic Richfield Co.

Introduction

The good old days were a long time ago. Now, though we must harness new technology and harsh climates to help provide needed energy supplies, we are also faced with the complex problem of satisfying not altogether consistent governments, the consumer, our banker, and someone's time schedule. Judging from the delays, massive capital overruns, and relatively low return this industry has experienced lately, it would seem that we have been missing something. At least one explanation is that we have not learned to deal with uncertainty successfully.

Some recent testing of SPE-AIME members and others gives rise to some possible conclusions:

1. A large number of technical people have little idea of what to do when uncertainty crosses their path. They are attempting to solve 1976 problems with 1956 methods.

2. Having no good quantitative idea of uncertainty, there is an almost universal tendency for people to understate it. Thus, they overestimate the precision of their own knowledge and contribute to decisions that later become subject to unwelcome surprises.

A solution to this problem involves some better understanding of how to treat uncertainties and a realization that our desire for preciseness in such an unpredictable world may be leading us astray.

Handling Uncertainty

Our schooling trained us well to handle the certainties of the world. The principles of mathematics and physics work. In Newton's day, force equaled mass times acceleration, and it still does. The physicists, when they found somewhat erratic behavior on the atomic and molecular level, were able to solve many problems using statistical mechanics. The extremely large number of items they dealt with allowed these probabilistic methods to predict behavior accurately.

So we have a dilemma. Our training teaches us to handle situations in which we can accurately predict the variables. If we cannot, then we know methods that will save us in the presence of large numbers. Many of our problems, however, have a one-time-only characteristic, and the variables almost defy prediction.

You may embark on a new project whose technology differs from that used on other projects. Or perhaps your task is to perform a familiar project in a harsh environment. Try to estimate the total cost and completion time. Hard! You cannot foresee everything. And, for some reason, that which you cannot foretell seems to bring forth more ill than good. Hence, the predictions we make are often very optimistic. Even though we see the whole process unfolding and see estimate after estimate turn out optimistic, our next estimate more than likely will be optimistic also.

What happens? Is there some deep psychological phenomenon that prevents our doing better? Because we are paid to know, do we find it difficult to admit we do not know? Or can we obtain salvation through knowledge? As we were trained to handle certainty, can we also find a better way to estimate our uncertainty?

I think so, but it will take some special effort — just as it did when we first learned whatever specialty that

What do you do when uncertainty crosses your path? Though it seems that we have been taught how to deal with a determinate world, recent testing indicates that many have not learned to handle uncertainty successfully. This paper describes the results of that testing and suggests a better way to treat the unknown.

got us into the business. As one of the Society's Distinguished Lecturers for 1974-75, I had a unique opportunity to collect information on the way our membership treats uncertainty. I do not claim that what you are about to read will set the scientific or business communities to quaking (others have noticed similar phenomena before[1]). But there are lessons that should help to improve our perceptions of uncertainty and, we hope, increase our economic efficiency by giving us better information on which to base decisions.

SPE-AIME Experiment

The experiment went like this. Each person put ranges around the answers to 10 questions, ranges that described his personal uncertainty. The questions were the following:

1. In what year was St. Augustine (now in Florida) established as a European settlement?
2. How many autos were registered in California in 1972?
3. What is the air distance from San Francisco to Hong Kong in miles?
4. How far is it from Los Angeles to New Orleans via major highways in miles?
5. What was the census estimate of U.S. population in 1900?
6. What is the span length of the Golden Gate Bridge in feet?
7. What is the area of Canada in square miles?
8. How long is the Amazon River in miles?
9. How many earth years does it take the planet Pluto to revolve around the sun?
10. The English epic poem "Beowulf" was composed in what year?

For some, the task was to put a 90-percent range around each answer. The person would think up a range such that he was 90-percent sure the range would encompass the true value. For example, in one section a gentleman put a range of 1500 to 1550 on Question 1. He was 90-percent sure that St. Augustine was established after 1500, but before 1550. In his view, there was only a 5-percent chance that the settlement came into being after 1550. If he were to apply such ranges for many questions, we would expect to find about 10 percent of the true answers outside of his intervals.

Other groups were asked to use 98-percent ranges — virtual certainty that their range would encompass the true value. I also asked for ranges of 80, 50, and 30 percent. The 30-percent interval would supposedly allow 70 percent of the true answers to fall outside the range.

Most sections used a single probability range. However, a few groups were divided in two, with each half using different intervals, usually 30 and 90 percent. I shall refer to these ranges as probability intervals.

You may want to test your skill on the test, too. The answers are in the Appendix. Use a 90-percent interval so you can compare with results given later.

Results and Conclusions

My testing turned up traits that should be of interest. [From this point on, the people referred to are the 1,200+ people at the local section meetings who answered the questions sufficiently to be counted. There were a significant number (350 or so) at the meetings who either had no idea of how to describe uncertainty or thought it chic not to play the game.]

1. People who are uncertain about answers to a question have almost no idea of the degree of their uncertainty. They cannot differentiate between a 30- and a 98-percent probability interval.
2. The more people know about a subject, the more likely they are to construct a large probability interval (that is, one that has a high chance of catching the truth), regardless of what kind of interval they have been asked to use. The converse seems to hold also: the less known, the smaller the chance that the interval will surround the truth.
3. People tend to be a lot prouder of their answers than they should be.
4. Even when people have been previously told that probability ranges tend to be too small, they cannot bring themselves to get their ranges wide enough, though they do somewhat better.
5. Simultaneously putting two ranges on the answers greatly improved performance, but still fell short of the goal.

Such conclusions come from the following observations. Looking at the data collected on each of the sections, we find that the average number of "missed" questions was close to 68 percent. We could adopt the following hypothesis:

> SPE-AIME sections will miss an average 68 percent of the questions, no matter what probability ranges they are asked for.

Mathematical statisticians have invented a way to test such hypotheses with what they call confidence intervals. They recognize, for instance, that the Hobbs Petroleum Section average of 6.26 misses out of 10 questions is subject to error. Slightly different questions, a different night, a longer or shorter bar — all kinds of things could conspire to change that number. By accounting for the variability of responses within the Hobbs chapter and the number of data points that make up the average, these statistical experts can put a range around the 6.26 much like the ranges the members were asked to use. Except that (unlike the members) when the statistician says he is using a 95-percent range, he really is!

For Hobbs, that range comes out to be 5.45 to 7.07. Since that range includes 6.8, or 68-percent misses, the statistician will agree that, based on his data, he would not quarrel with the hypothesis as it applies to Hobbs.

Table 1 shows all the 95-percent ranges and Fig. 1 illustrates how these ranges compare with the 68-percent hypothesis. You will see a portion of the Los Angeles Basin Section whose confidence interval (5.24 to 6.68) does not include 6.8. There are three possible explanations:

1. The group has a bit more skill at handling such a problem than most.
2. Being part of an audience that was asked to use two different ranges, there was a more conscious effort on their part to use a wider range.
3. The statistics are misleading, and the group is not different from the others. We expect this to happen about 5 percent of the time. (Our testing mechanism

TABLE 1 — SUMMARY OF 95-PERCENT RANGES

SPE-AIME Section	Number of Usable Responses	Requested Range (percent)	Expected Number of Misses	Actual Number Average Misses	95-Percent Confidence Interval
Hobbs Petroleum	34	98	0.2	6.26	5.45 to 7.07
Oklahoma City	111	98	0.2	7.00	6.64 to 7.36
Los Angeles Basin (1)	28	90	1	5.96	5.24 to 6.68
San Francisco	61	90	1	6.41	5.89 to 6.93
Oxnard	26	90	1	7.38	6.64 to 8.12
Long Beach (1)	28	90	1	6.04	5.20 to 6.88
New York	29	90	1	6.52	5.76 to 7.28
Bridgeport/Charleston (1)	16	90	1	7.63	6.89 to 8.37
Anchorage	63	90	1	6.54	6.00 to 7.08
Bartlesville	44	90	1	6.30	5.61 to 6.99
Lafayette	79	90	1	6.51	6.03 to 6.99
Shreveport	41	90	1	6.83	6.18 to 7.48
Vernal	13	80	2	7.23	6.30 to 8.16
Denver	129	80	2	6.46	6.12 to 6.80
Cody	42	80	2	7.31	6.74 to 7.88
Columbus	27	50	5	6.96	6.47 to 7.45
Lansing	30	50	5	6.83	6.16 to 7.50
Chicago	41	50	5	6.54	5.97 to 7.11
Tulsa	53	50	5	6.79	6.33 to 7.25
Los Angeles Basin (2)	27	30	7	7.00	6.26 to 7.74
Long Beach (2)	28	30	7	7.39	6.80 to 7.98
Bridgeport/Charleston (2)	15	30	7	7.82	6.97 to 8.67

was a 95-percent confidence interval.)

Likewise, the Bridgeport/Charleston (W. Va.) sections had ranges that did not encompass 6.8. In their defense, the meal service had been poor, the public address system had disappeared, and there were more than the normal misunderstandings. Even so, their lower limits of 6.87 and 6.97 just barely missed the 6.8 target.

One group of highly quantitative people also took the test. I mention this group because of the large number of members it includes and because it provides evidence that the more quantitative people may do a little better in estimating uncertainty — but still not as well as they would like. (See Table 2.)

The 68 percent would not be expected to hold on all kinds of questions or all kinds of people. In fact, it is clear that the number would have been higher had it not been for relatively easy questions such as Questions 1 and 4. Most people know St. Augustine was a Spanish community and, therefore, had to be established between 1492 and 1776. By making the range a bit more narrow than that, they could be reasonably sure of bracketing the true answer. Even so, more than one-third of the members missed that one — regardless of their instructions on range.

Based on a sample of the 1,200+ quizzes, here are the average misses for each question:

Question	Average Misses (percent)
1	39
2	67
3	60
4	50
5	69
6	68
7	76
8	69
9	74
10	85

Questions such as Questions 9 and 10 were difficult, and we found 80 percent or so misses — again regardless of the requested probability of a miss.

People who have no idea of the answer to a question will apparently try to fake it rather than use a range that truly reflects their lack of knowledge. This trait may be as universal a part of human nature as laughter; certainly it is not peculiar to SPE-AIME members.

Is the Problem Costly?

Why should anyone get excited about such results? Because, I think, similar behavior on the job can cost industry a bundle. Our membership at various levels of-

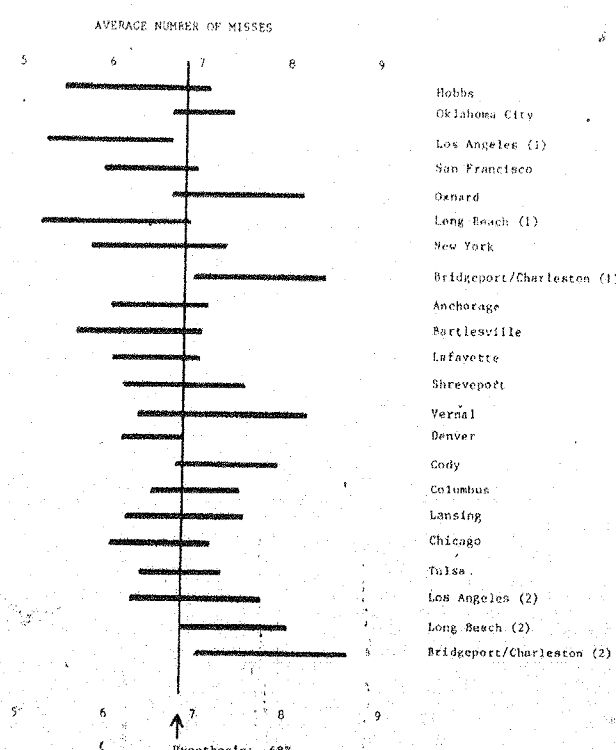

Fig. 1 — The 95-percent confidence intervals of SPE-AIME sections. Average number of misses on 10-question quiz.