

The Good Judgment Project

Hol White

Manufacturing equipment makers host a Machine tool show at McCormick's Place in Chicago every other year in September. Machine tool vendors come from around the world to demonstrate the latest in machining technology. My visit to Chicago however is not all work and my family loves to go with me. We often stay at a hotel downtown, not far from the John Hancock Building. My wife and daughters enjoy shopping on Michigan Avenue; the ladies have their own personal assistant at Nordstrom's that helps guide them to the most absolutely wonderful shoe that Kate Spade makes. I on the other hand will go to the show with my dad.

This year, the show's two top attractions were Fanuc's robotics and 3d printing. The robots on display were straight out of a science fiction movie. The machines had greater dexterity than I had ever seen before. The Robots were assembling fuel valves, all without a break, lunch or need to register a complaint. Across the hall metal parts were printed on a printer, complete, without the need for multiple machines. Was creative destruction of manufacturing on the way? There was a lot more to see but after a full day on my feet the Polish Dog I had for lunch was no longer sustaining me. I was ready to get back to the Whitehall hotel and go out for dinner.

When I meet back up with the girls, they show me their purchases; we have a glass of wine, then head down stairs to get a cab. Our door man flags our ride. Our destination is Topo Gigios, an Italian restaurant in old town that has wonderful calamari. A Chicago cab is a tight squeeze for a family of five so I ride in front. When we come to the stop I glance down and can tell from the Cabbie's name on his license that he may not be from Chicago. I ask where he is from. He says that he emigrated here from Nigeria six years ago, ahh I say, that is a long way from home.

Nigeria, I think to myself, wasn't there a question I had about Nigeria? This man may have information I am looking for. I proceed to ask, "In your opinion will the government of Nigeria begin official talks with Boko Haram prior to December 31, 2012?"

Our driver grimaces, picks up speed and becomes less patient with traffic. At the next light he turns to me and says, do you know what Boko Haram means? It means "western education is sinful". The government of Nigeria should never negotiate with terrorist groups. The Boko Haram are bad people. He goes onto explain that he was a Nigerian police officer and had to leave. Our driver then executes a sudden lane change around a stopped Chicago Police car and mutters something, the joviality in the back seat ceases as it is clear I have touched a sore spot with our driver. We pull up to the restaurant. I open the cab door leaving any follow up questions I have in the cab and with my fare and a generous tip. My family exits the Cab quickly and turns to ask me why on earth was I asking a Chicago cab driver about the Boko Haram in Nigeria?

Mr. President, Mr. Secretary, fellow presenter and members of the Athenaeum, tonight's paper explains the reason for the Boko Haram question, it is about The Good Judgment Project. The paper will provide background on a World event forecasting tournament sponsored by the US. The paper will also describe how the Good Judgment project plans to excel in that tournament, and as a special treat there will also be a simple quiz designed to reduce your confidence. I will proceed to share two example questions and with my answers and rationale. I will conclude with why I participate in the project.

The need for the Good Judgment Project became clear on a September morning as I was driving to work. Cars ahead of me were pulling off to the side of the road. I rolled down my window and could not hear a siren nor see the lights of any emergency vehicle. On the car radio there was a special report out of New York that a plane had crashed in to the World Trade Center.

At that moment and for the days that followed 9/11 there was a feeling of helplessness followed by thoughts that we Americans should have known and perhaps we did know but failed to act. Despite massive spending on gathering and analyzing intelligence, the government of the United States was unable to prevent the World Trade Center attack.

One of the government's early attempts at developing better methods for anticipating attacks included creating a betting pool. A pool where participants could bet on the possibility of a terrorist event occurring. Congress however was wary of putting up a betting window that might offer perverse incentives. The bettors might be tempted to help their bets along.

The need for identifying potential threats was still there. Was there a way the government could integrate the collective knowledge of all the armed services and intelligence agencies? How could they connect the dots prior to any future attacks?

In 2004 Congress created the Office of the Director of National Intelligence. The act gave the Director the task of bringing together the government's 17 intelligence organizations. It was clear that research in new methods was needed because there was very little sharing of information between the agencies. The Director created the Intelligence Advanced Research Project Agency with the mission to invest in high risk, high payoff research programs that have the potential to provide our nation with overwhelming intelligence advantages over future adversaries.

IARPA funded a four year, 50 million dollar study to predict major world events including wars, regime changes and disease outbreaks. Unlike the earlier program where betting was allowed participants could not profit from their predictions. The study would provide precisely worded questions to elicit clear responses to enable measurement of forecast accuracy. The results of each question were to be settled by independent confirmation of news source such as the Economist, BBC or Reuters.

The study was designed to determine whether the intelligence community could aggregate the judgment of thousands of analysts rather than rely on the expertise of just a few and provide more accurate warning to policy makers before a major global event. The study also solicited and accepted proposals from five teams outside the intelligence community as competition.

Aggregation theory also known as the Wisdom of the Crowds was going to be used to reduce forecasting error. The theory was discovered by an English polymath/ Renaissance man named Galton who liked to measure things. Galton had gone to a county fair where a contest was being held to guess the weight of a steer once butchered. Galton wrote down all the guesses entered by the fair goers. He was surprised to learn that when he averaged the fair goers' guesses it was within a pound of the actual measurement. The average was also better than the butcher's and cattlemen's guesses.

The University of Pennsylvania and the University of Berkeley submitted a joint proposal and which was accepted by IARPA into the tournament and The Good Judgment Project was formed. Leading the project is Phillip Tetlock, Barbara Meller and Don Moore. Professor Tetlock, author of "Expert Political Judgment" maintains that experts are no better at making predictions on election results than you are me. The two other project leads of Meller and Moore are experts in decision making and overconfidence respectively.

I joined the Good Judgment team after reading a blog at Marginal Revolution authored by Tyler Cowen, an economist that I follow. The entry on the blog to join was the following;

Phillip Tetlock requests your Help

He is one of the most important social scientists working today and he requested that I post this appeal:

Starting in mid-2011, five teams will compete in a U.S.-government-sponsored forecasting tournament. Each team will develop its own tools for harnessing and improving collective intelligence and will be judged on how well its forecasters predict major trends and events around the world over the next four years.

If you're willing to experiment with ways to improve your forecasting ability and if being part of cutting-edge scientific research appeals to you, then we want your help.

Wow, the idea intrigued me and I followed the link provided to the website and signed up ^{for} the good judgement project

The social scientist Tetlock saw that anyone claiming to be an expert in the subject tended to be overconfident in their predictions. Meller and Moore thought that a group of lay forecasters could be taught to minimize overconfidence. Special training was to be given to the forecasters including myself to learn to frame their answers by thinking stochastically (meaning ^{our} their answers include accommodation for random variables and chance) rather than providing deterministic solutions (where everything that happens | results from having occurred before). In horse race parlance, there needed to be more betting on the horse to show rather than to win, even if the horse had won in the past.

To demonstrate the training of the process of providing a range of answers with a confidence level let's now do the quiz you cannot fail, I have provided pencils and ^{quiz} paper at all the tables for your use. I

^{am} will give ^{us} you 10 questions to answer. You need to give your answers as a range that you feel 90% confident the answer will fall in that range. If the question was: How long will next month's Athenaeum society be? My range answer would be I am 90% sure that the next Athenaeum Society Meeting will be between 1 hour and 1 ¼ hour long. However that is not your question, here are your questions.

1. Amount of Time elapsed since the big bang? 13.8 billion years
2. Year that John Steinbeck won the Nobel Prize in Literature? 1962

- | | |
|---|--|
| 3. Distance to the moon? | 239,000 miles, 380,00 km |
| 4. Air Distance from Madrid to Baghdad? | 2,700 miles, 4,300 km |
| 5. Year that the construction of the Roman Coliseum was finished? | 80 AD |
| 6. Year Magellan circumnavigated the globe? | 1522 |
| 7. Height of the Aswan dam? | 375 ft., 114 meters |
| 8. Year Ghandi was born? | 1869 |
| 9. Surface Area of the Mediterranean Sea? | 970,00 square miles, 2.5 million square km |
| 10. Gestation Period of the great blue whale? | 335 Days |

Note: Review the Answers above to the right.

Was there any one that got all ten questions where the answers were within their range? Good smart man. How about 9, even better these individuals are perfectly calibrated. How about 8, 7, 6?

The Good Judgment Project used training like these questions above to help Forecasters calibrate their answers and not be too overconfident.

With participants selected and trained The Good Judgment Project was ready for the tournament. IARPA submitted a list of a 10 questions to each of the teams to forecast every couple of weeks. The Good Judgment Team posts the questions on a website and provides a market (similar to ^{five} a trade) for the forecasters to ^{buy into the} give their likelihood of whether or not the event will occur in terms of a percentage. ~~Aggregation occurs as~~ ^F forecaster assign probabilities to the likelihood of an event occurring by buying shares, the higher the likelihood the forecaster believes the event will happen the higher the price of the shares. Conversely one can sell shares if one thinks the event is less likely to ^{own} have than the forecast.

Here are two of those questions IARPA has given and how I think about them.

Question 1: Will the Free Syrian Army gain control of Aleppo by October 1?

In July of this year the other forecasters were predicting that the Syrian city of Aleppo would fall. There were several good reasons to think that the fall was to occur soon. Our own Secretary of State Hillary Clinton was pushing for Assad to step down and the Free Syrian Army was having success taking down Syrian border Crossings.

I thought Assad's army losing control of Aleppo was not likely for in the short term for four reasons. One there were reports that the cost of ammo was escalating, the Free Syrian Army was going to have to conserve on resources 2) taking and controlling a city was entirely different than overrunning a rural border post, 3) the Free Syrian army had lost the city populations support after the summary execution of civilians that had supported the Assad government and 4) The city of Aleppo holds the Citadel a historic site that holds significance for the army and Assad's regime. I estimated that there was between a 15 and 25% chance that the City of Aleppo would fall to the Free Syrian army and wagered heavily when other Forecaster predicted over 80%

By October 1 the Free Syrians Army had lost momentum as their push into cities had begun to stall, my forecast was correct.

Question 2: Otherwise known as the Wild Chicago Taxi Ride question. Will the Nigerian Government hold official talks with Boko Haram by December 31, 2012?

Where the media had full coverage of the Syrian Civil War there is not a lot of information coming out of Nigeria on possible talks with Boko Haram. I only had a gut feeling that there was a greater than 50% chance of the Government might be having talks with the group. This is the reassuring confidence level of flipping a coin. The only supporting information was that there was an African news

report of a Nigerian official stating the kitchen table talks were happening with the group. On that basis I ^{bought} placed a small amount on the forecast, ^{of shares} at 50%.

That was my thought on the question prior the ride with the cabbie in Chicago, noting his strong reaction and rebuke of Boko Haram made me investigate the question further. The last leader of Boko Haram was killed while in police custody several years ago, the killing was an act of retribution for lives of police officers taken by Boko Haram. There was currently no apparent leader of Boko Haram so government discussions with the group are probably pointless. My 50% forecast has fallen to the group's consensus of 10% and my ^{shares} wager took a hit.

The first round of the tournament contained many similar questions that the group had to answer. At the end of the first round the Good Judgment Project led all other groups in the tournament. Professor Tetlock said it was remarkable to find a government agency willing to test their best and brightest against those outside the system.

For me there are three main benefits for me participating in the project.

The first benefit is meeting a small bit of patriotic duty from helping our nation find better ways to utilize our intelligence resources and addressing the sense of helplessness felt after 9/11.

The second benefit is that the question ⁵ provides a measure of my ability to learn about new situations in other parts of the world and make predictions. The news stories I read now, require a more in depth understanding so they can be set in context of the questions that IARPA provides.

The third benefit is that the study allows me opportunities to connect with others I meet, whether it be the son of a neighbor who studied a year in Egypt (asking him about who the next president of Egypt will be, with the response was that Muslim brotherhood's candidate didn't stand a chance), my son's fraternity brother who is from Ghana (asking who he thinks the next president of

Ghana will be with his response being that his granddad knows both of the candidates and says either one of them would be good and last but not least a Taxi driver from Chicago.

Donald Rumsfeld once said, there are known knowns. These are the things we know that we know. There are known unknowns. That is to say there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know that we don't know.

Forecasts regarding what might happen in the future can never be reduced to a certainty. It is thinking about the issue, placing it in context and having an understanding of our own over confidence that allows us to give a best qualified answer. The Government's sponsored forecasting tournament and the Good Judgment Project are providing insights to the process that may help protect our nation from attack. The insights from the project could also be used to anticipate humanitarian aid in areas of conflict reducing human suffering. In business I use my understanding of overconfidence to help poke holes in sales forecasts under review or lower projected cost savings.

In conclusion: Of the known knowns, I know my gals will find some Kate Spade shoes at Nordstrom's in Chicago they love, I know I don't know how much robots and 3D printing will affect manufacturing in the long term and finally I don't know what I don't know.

Thank you for you participation and you^v interest in this topic and for participating in my short quiz.