Journal of
Professional Communication jpc

Slow down and smell the data

Andrew Laing★

Cormex Research, Toronto (Canada)

A R T I C L E I N F O	**A B S T R A C T**
	This review of Nate Silver's *The Signal and the Noise* discusses how Silver's celebrity statistician status has been achieved and is well-earned. He touches on how the book reinforces, in an understandable fashion, that a conceptual understanding of Big Data is a crucial element of successful professional communication practice. This positive review identifies the strengths of Silver's book – namely its solid organization and its use of real-world examples. *The Signal and the Noise* is recommended for professional communicators who want to gain an understanding of data and the many uncertainties inherent to its analysis.

Nate Silver has accomplished a rare Einstein-like feat: he's a celebrity statistician. As the former *New York Times* author of the popular *FiveThirtyEight* blog – he recently parachuted into the Disney family of ESPN and ABC properties to provide sports and political commentary and the blog will be going with him. Silver became a media star of the 2012 US presidential election campaign; interviews on numerous media outlets, including *The Daily Show with Jon Stewart* and *The Colbert Report* followed after he successfully predicted the outcome of the presidential election in all 50 states as well as 31 of 33 Senate races, replicating a feat he'd achieved in the 2008 presidential race. And in his spare time in 2012, Silver published his book, *The Signal and the Noise: Why So Many Predictions Fail – But Some Don't.*

Silver's fame is not a reason to read this book but there are two other reasons that should prompt anybody, particularly professional communicators, to pick it up. The first is that it is a refreshing and enlightening tonic to the relatively bad news around political polling. Political polls in Canada are now 0-for-3 in providing any reasonably accurate prediction of provincial elections, and

★*Corresponding author (Andrew Laing)*
Email: andrew.laing@cormex.com

pollsters have been taking increasing heat for this. Silver's *FiveThirtyEight* is based on a unique approach to weighting and aggregating political polling data and, in explaining this process in *The Signal and the Noise*, readers of political polling will get a much better understanding of the limitations of polling beyond the current discussion over sampling methods that dominate media discussion of the issue. The second issue is Big Data. Big Data is now a *big deal*, and the hype and counter-hype is a recipe for a poor understanding of what it entails. The premise of Silver's book squarely addresses the challenge of understanding Big Data at a conceptual level – that throughout history humanity has been confronted by 'leaps' in the production of information that has not always been matched by our ability to understand and use it, and that Big Data is simply another episode.

The book is laid out in a series of chapters, each dedicated to an event or issue that a reader can immediately relate to. Each chapter quickly gets down to examining the collision of prediction and information, including the 2008 collapse of the US housing market and its effects on the economy, tornado and weather prediction, earthquake analysis, H1N1 and pandemic modeling, the *Moneyball* world of predicting professional sports team performances, terrorism and counter-terrorism, gambling, climate change and, of course, polling and election outcomes.

Through each example, Silver introduces key concepts about prediction and explains the high failure rate of many models designed to predict outcomes. Readers of Malcolm Gladwell will recognize and appreciate Silver's conversational style and his prodigious use of interviews with experts and colourful, off-the-radar characters to make his points. In fact, the biggest strength of the book may be its accessibility. Statistical points are illustrated using easy-to-follow line charts, scatter plot graphs and other devices. For those worried about the stats and reliving the nightmare of their university research methods courses, there are no calculations save for that of Bayes's Theorem (see below). The usual Statistics 101 concepts such as how to calculate a correctly estimate a p-value are entirely absent.

Halfway through the book, after introducing readers to a number of concepts and limitations behind predictions, Silver switches gears and introduces Bayesian theory – a key element to the remaining chapters in which he attempts to shift readers into thinking about the world probabilistically. Bayes's Theorem postulates that the probability behind a hypothesis is based on three determinants:

- a probability of a condition of the hypothesis being *true*;
- the probability of a condition of the hypothesis being *false*; and

- a *prior* estimate of the probability of the event before it actually occurs.

True to the approachable nature of the book, Silver outlines Bayes's Theorem using the whimsical example of a woman estimating the probability that her husband is having an affair by finding a pair of panties in his drawer (for the record, it is 29%), but he then applies it to more serious statistical probabilities people encounter: why many oncologists recommend that women not get a mammogram until their 50s, and why US intelligent agencies did not predict the attacks of 9/11. More importantly, Silver also reviews how many modern statistical models are based on R.A. Fisher's sampling approach that rejected the subjective bias of Bayesian *priors* in favour of "frequentism" – increasing the predictability of a model by reducing the chance of error through repeated samplings. Silver's juxtaposition of Bayes and Fisher is a highlight of the book.

The wide range of real-life examples and empirical data sets that Silver employs, while a strength, is also at times a weakness. In some areas a reader may want further detail to be convinced of Silver's point, or take exception to some of his conclusions. For example, in the chapter devoted mostly to looking at infectious disease surveillance modeling involving pandemic influenza viruses such as H1N1, Silver goes on a tangent to look at other areas of public health, such as HIV infection rates and reported cases of Autism Spectrum Disorder. According to Silver and relying on the conclusions of a researcher at the Harvard School for Public Health, the almost one-to-one correlation between the rising volume of media coverage of autism and the number of diagnoses of the disease in the US over a sixteen-year period suggests that public awareness to autism through media reporting may explain the rising number of diagnoses offered – a contentious conclusion to say the least. However, these instances are found very rarely in the book and do not take away from his main arguments.

Silver's conclusion – the need to embrace Bayesian uncertainty and complexity in data – is a welcome conclusion, but one that many professional communicators may find discomfiting. Communicators often have a *yin-yang* relationship with empirical data and measures: they want a credible, communicable means of measuring results, yet they come to the game with little understanding of how to work with data, and tend to bias towards the simpler, grander statements that they believe better communicate performance objectives. Everyone in business – including professional communicators exposed to content made available through digital and social media channels – faces the opportunities and challenges of Big Data. Embracing the complexity, the probabilistic nature, the messiness or "noisiness" of data, and being willing to

test and retest hypotheses and "slow down and smell the data" as Silver wryly puts it, are the main lessons professional communicators can learn from this engaging book.

References

Silver, N. (2012). *The signal and the noise: Why so many predictions fail – but some don't.* New York, NY: Penguin Press.