

This article first appeared in the January 2007 edition of the Quarterly Member Newsletter of International Association for Information & Data Quality. The IAIDQ is the world's leading not-for-profit Professional Association for Information Quality professionals, founded in 2004 and engaged in taking the profession to greater heights through an industry wide, vendor-neutral certification program. Be a part of the IQ revolution. Join us. Membership benefits include access to a treasure of Data Quality related knowledge base, access to thought leadership from multiple Data Quality experts in the form of thought provoking articles, webinars, book summaries and opportunities to engage in lively professional discussions. Full details about the Association and what they do can be found on the [IAIDQ website](#).

## The Freakish World of Root Cause Analysis

By  
Daragh O'Brien

### INTRODUCTION

On a recent project I had to guide a stakeholder through a root cause analysis activity to determine the real root causes of an Information Quality Problem.

My stakeholder had a number of assumptions that had been paraded as fact for some time and as such, everyone believed that tackling these “known” issues would deliver an antidote to the business problem that was the focus of the project. So strong was the stakeholder’s confidence in these assumptions that he believed we could dispense with our scoping and definition of metrics for the quality of the process and the identification of the various scenarios that existed in the raw data. All we needed to do was prove that native use of a legacy system, or a bug in a call-centre system, was the problem and get ‘The Business’ to sort it out.

Unfortunately, the data simply did not support that analysis. Our own preliminary study, based on business rules for the flow of data through the various stages of the process, suggested that the actual root cause lay elsewhere, with native use of systems and other causes being lesser contributors. Unfortunately, our early findings flew in the face of the conventional wisdom so it was a hard sell. Our analysis must be wrong. Recheck the data. Shoot the messenger.

“Conventional wisdom” is a phrase that was coined by the economist J.K Galbraith<sup>1</sup> to describe ideas that are generally accepted as true by society. He did not intend it to be a compliment. Indeed, many urban legends are accepted as true on the basis of conventional wisdom. Statements which are constantly repeated become conventional wisdom regardless of their truth or accuracy, because “audiences of all kinds most applaud what they like best.”

Galbraith wrote that “*the enemy of conventional wisdom is not ideas but the march of events*”. More recently, in *Freakonomics*<sup>2</sup>, Steven D. Levitt and Stephen J. Dubner have highlighted the importance of rigorous analysis of data to identify significant root causes for key outcomes<sup>3</sup>. That last sentence sounds a lot like Kaoru Ishikawa’s admonition to ‘speak with data.’ More importantly, all their advice flies in the face of ‘conventional wisdom.’

---

1 The phrase was coined by Galbraith in his book *The Affluent Society* in 1958.

2 *Freakonomics: A rogue economist explores the hidden side of everything*, Levitt, Steven D. & Dubner, Stephen, Penguin Books, 2006 (refers to paper back edition)

3 As an aside, note the potential Information Quality problem here with the given names of the authors... two different spellings of the same name.

*Freakonomics* is full of examples of how analysis of the data and pursuit of the root cause yields answers that are very different from that provided by conventional wisdom.

For example, conventional wisdom attributed the non-emergence of a massive wave of violent crime in the US to improved gun laws, zero tolerance policing and efforts to reduce poverty.

These may have been contributing factors, but according to Levitt and Dubner a more significant causal factor<sup>4</sup> is the landmark U.S. Supreme Court case *Roe v. Wade*. Simply put, the change in the law relating to abortion in the 1970s resulted in many of the criminals that were forecast to be running around in the 1990s simply not being born; ergo, no massive increase in violent crime. The improvements in gun control laws, improved policing and campaigns to reduce the impact of poverty may have contributed further to avoiding the predicted wave of violent crime, but these may not have been as key a factor as conventional wisdom might lead us to believe.

Levitt's analyses and Dubner's prose are persuasive. Their arguments are made even more compelling by the fact that they have actually *looked at the data* to determine if there is a correlation. Irrespective of one's opinion on the causal factor (it *is* an issue fraught with moral and religious debate) the data presented in *Freakonomics* is stark, with a number of strong correlations cited.

The above example is an important lesson about 'speaking with data.' Often what the data is telling you is not what people might want to hear. Often it flies in the face of conventional wisdom. As Levitt and Dubner tell us in chapter 5 of the book, "*Separating facts from rumours is always hard work.*"

Colleagues of mine at an Irish Blog called [Tuppenceworth.ie](http://tuppenceworth.ie)<sup>5</sup> have recently 'spoken with data' about the state of the Irish print media. In effect, their blogpost was a rough IQ audit on Irish print media. They found a substantial amount of 'advertorial' or press-release regurgitation passing as journalism, with real journalism being in a stark minority overall. This finding flew in the face of the conventional wisdom that journalists are part of a 'fourth estate,' and led to a number of comments on several blogs, and an article in a leading Irish daily newspaper which seemed to miss the point of the study entirely<sup>6</sup>.

The survey was not entirely scientific and would certainly benefit from a more rigorous approach (the instigators will admit this without challenge). However, the facts presented are persuasive and prompted some media outlets to pick up on the findings, with some journalists taking the study as a wake up call<sup>7</sup> and the Secretary of the Irish National Union of Journalists admitting at a conference in early December<sup>8</sup> that standards in journalism had indeed slipped. This admission was, of course, unrelated to the findings of the PaperRound study.

---

4 An extract from this chapter can be read at <http://www.freakonomics.com/ch4.php>

5 The blogpost is entitled *Paper Round – Weekly Herald* and can be read at the Tuppenceworth.ie blog at <http://www.tuppenceworth.ie/blog/index.php/2006/12/11/weekend-herald>. See also [www.tuppenceworth.ie/paperroundwiki](http://www.tuppenceworth.ie/paperroundwiki)

6 A scanned copy of this article can be found at [http://www.flickr.com/photos/editor\\_tupp/305644595/](http://www.flickr.com/photos/editor_tupp/305644595/)

7 Example, Shane Hegarty in the Irish Times, Saturday 3rd December, text is accessible at <http://www.thestory.ie/2006/12/04/can-blogs-beat-papers-to-a-pulp/>

8 See post on <http://www.tuppenceworth.ie/blog/index.php/2006/12/05/factbobs>

*Such is the power of speaking with data.* Everyone knew the state of our papers, the Tuppenceworth project merely quantified it. Readers interested in learning more can check out the blog and wiki referenced in the footnotes.

So how did I win my project stakeholder over to the reality of the root causes?

- I asked him to review and revise with me the business rules that underpinned the analysis we had conducted. I did this to reinforce the ownership he had over those rules (it was his process after all). [**Define the data**]
- I asked him to review the design and test collateral of the analysis process (we'd developed it for him after all), essentially re-running the testing phase of the first part of our project. [**Validate analysis and analysis data**]
- I helped him design a series of tests to run on the call centre system and on the 'native' legacy system to check for compliance with business rules at each point in the process. My team also helped run the tests as we had extensive experience in testing software in a previous incarnation. [**Gather fresh data**]

Once my stakeholder accepted that the analysis and business rules were correct, he could no longer reject the idea that conventional wisdom might be wrong. From that point, the gathering of new data by running the test cases meant we could identify points of failure in the 'live' process and validate the root causes that the original analysis had suggested.

So, did conventional wisdom prevail? Much like the crime wave which failed to engulf America in the 1990s, the causes identified by conventional wisdom were still contributing factors. Native use of a legacy order/billing system *was* a contributing factor. However, the stakeholder's firmly held view that there was a bug in the call centre system that handled sales and order processes was not supported by any of the data, so we avoided wasting time tracking down a phantom. The actual significant root cause was something that (almost) none of us had ever considered. It flew in the face of conventional wisdom. But the data supported the finding.

Your Information Quality projects may throw up similar challenges. I'd be interested in hearing your approaches to getting your stakeholders to 'hear the data.'

---

### About the Author



**Daragh O'Brien** Taoiseach (Chief Executive) of Castlebridge Associates, an Information Quality Consulting and Mentoring practice based in Wexford, Ireland. Prior to establishing Castlebridge Associates, Daragh spent 12 years with a large Telecommunications company in a variety of challenging Information Quality related roles ranging from CRM implementations (including Single View of Customer), to the delivery and management of Compliance and Governance functions.

He is a Charter Member of the IAIDQ, a Fellow of the Irish Computer Society, and served on the IAIDQ's Board of Directors from 2005 to 2009. He is a frequent speaker at conferences worldwide, and has written countless articles on Information Quality topics, as well as maintaining his personal blog ([obriend.info](http://obriend.info)) and the IAIDQ's IQTrainwrecks blog ([iqtrainwrecks.com](http://iqtrainwrecks.com)). Daragh has also devised course modules in Information Quality and related disciplines for FETAC ([Fetac.ie](http://Fetac.ie)) and Dublin City University ([DCU.ie](http://DCU.ie)).

He can be reached via email at [daragh.obrien @ iaidq.org](mailto:daragh.obrien@iaidq.org)

© 2006 Daragh O'Brien