sportsology

RESEARCH ACADEMY

BRIDGING THE INTERPRETER GAP: THE ANSWER TO THE BIG DATA DILEMMA *OR* WHERE'S C-3PO WHEN YOU NEED HIM?

BRIDGING THE INTERPRETER GAP: THE ANSWER TO THE BIG DATA DILEMMA *OR* WHERE'S C-3PO WHEN YOU NEED HIM?

Professor Chris Brady

Despite the hype around the value of big data, much of the research and anecdotal evidence suggests that the way in which we work and consequently the value added by big data has and will not change radically. Five years ago, the MIT Sloan Management Review (SMR) published a condensed version of this article.[1] In the last 6 months we have revisited the issue only to find that the problems identified in the original research remain stubbornly in place.

In an SMR article from nearly a decade ago the problem was clearly outlined –

> *according to many executives, the new crop of data scientists may be falling into the old traps. A common complaint is that data scientists are aloof and seem uninterested in the professional lives and business problems of less-technical coworkers; they don't see a need to explain or talk about the implications of their insights, which makes it difficult for them to partner effectively with professionals whose business expertise lies outside of the technical realm.*[2]

While Tom Davenport argued that the future belongs to the 'data scientists' it is also clear that the mythical creature he described is just that, mythical. The skill sets Davenport attributed to data scientists were truly magical. He said that data scientists,

> *understand analytics, but they also are well versed in IT, often having advanced degrees in computer science, computational physics or biology - or network-oriented social sciences. Their upgraded data management skill set — including programming, mathematical and statistical skills, as well as business acumen and the ability to communicate effectively with decision-makers — goes well beyond what was necessary for data analysts in the past.*[3]

However, even he admits that, *"this combination of skills, valuable as it is, is in very short supply"*. Anjul Bhambhri, vice president at Adobe, embellished Davenport's description even further when she stated that, *"A data scientist is somebody who is.....almost like a Renaissance individual"*.[4] As far as we can tell, only C-3PO, who boasts of being fluent in over 6 million forms of communication could fit the bill. C-3PO is the humanoid robot from the

[1] *Why Your Company Needs Data Translators*, MIT Sloan Management Review Winter 2017 edition.
[2] J. G. Harris and V. Mehrotra, "Getting Value From Your Data Scientists," MIT Sloan Management Review Magazine (Fall 2014).
[3] T. H. Davenport, P. Barth and R. Bean, "How 'Big Data' Is Different", Sloan Magazine: Fall 2012,
[4] IBM Institute for Business value, "Analytics: The new path to value: How the smartest organizations are embedding analytics to transform insights into action", Executive Report 2010.

Star Wars movies whose primary function is to assist with issues of customs and translations to enable ventures between different cultures to run smoothly.

More recent research and anecdotal evidence collected by the Sportsology research academy suggest that not much has changed. A key issue that continues to present is a recognition that there exists a missing link in performance management practice and, indeed, in the literature, between what has been termed 'big data' and the predominantly dismissive attitude of the decision-makers (GMs, head coaches, CEOs, COOs, etc.) to that data and those responsible for delivering it. Bridging that cultural gap could provide considerable competitive advantage to any organization concerned with high performance.

At various MIT Sports Analytics conferences, two of the most common themes were how to influence senior management and how to balance gut and data. As Ken Cantanella, Assistant General Manager at Sacramento Kings put it – *"analysts are trying to make an impact on the organization by transferring their knowledge to somebody that will actually make use of it"*. In a 2014 report, Tom Davenport cited English soccer club Manchester City as being, *"relatively advanced in terms of performance analysis"*[5]. Although Manchester City was among the leaders in developing an analytics department, that department's impact was significantly hampered by at least one manager who, at best, only tolerated the analysts, preferring to rely upon his own intuition as the predominant factor in his decision making.

While the recurring theme in this space has been the inability of senior management to understand or even accept the value of data-driven decisions in business and sport, no real solution has been suggested let alone implemented. The dominant argument has been that "data experts" will become more and more important. This issue is, therefore, a central performance management issue. The problem as we at Sportsology see it, is that both sport and businesses are asking the wrong people to step into this role and that the mythical beast expected to step into the role cannot exist. As a result of our research meetings, we accept that there is a significant gap between the data guys and the decision-makers, a gap that we want to call "the interpretation gap". Those who will eventually fill the gap we call 'data translators'.[6]

Where Davenport believes that data scientists can bridge the gap we believe that the role can only be filled by domain experts (coaches, COOs, business managers etc.). To date, business/sport has been trying to teach the quants guys (often recent graduates) about the business/sport in which they operate; it should be the other way around. It is easier for domain experts, with deep knowledge of the game/business in which they are engaged and, just as importantly, the requisite interpersonal skills, to learn about data than for data scientists to learn about the domain, especially the language of that domain. Domain expertise requires a high level of practical experience that it is difficult to acquire on a theoretical basis and it also lends itself more readily to the 'story-telling ability' that must be an essential skill requirement of the 'translators'. As Davenport admits,

> *data scientists [need to be able] to communicate in language that all their stakeholders understand - and to demonstrate the special skills involved in storytelling with data, whether verbally, visually, or - ideally – both.*[7]

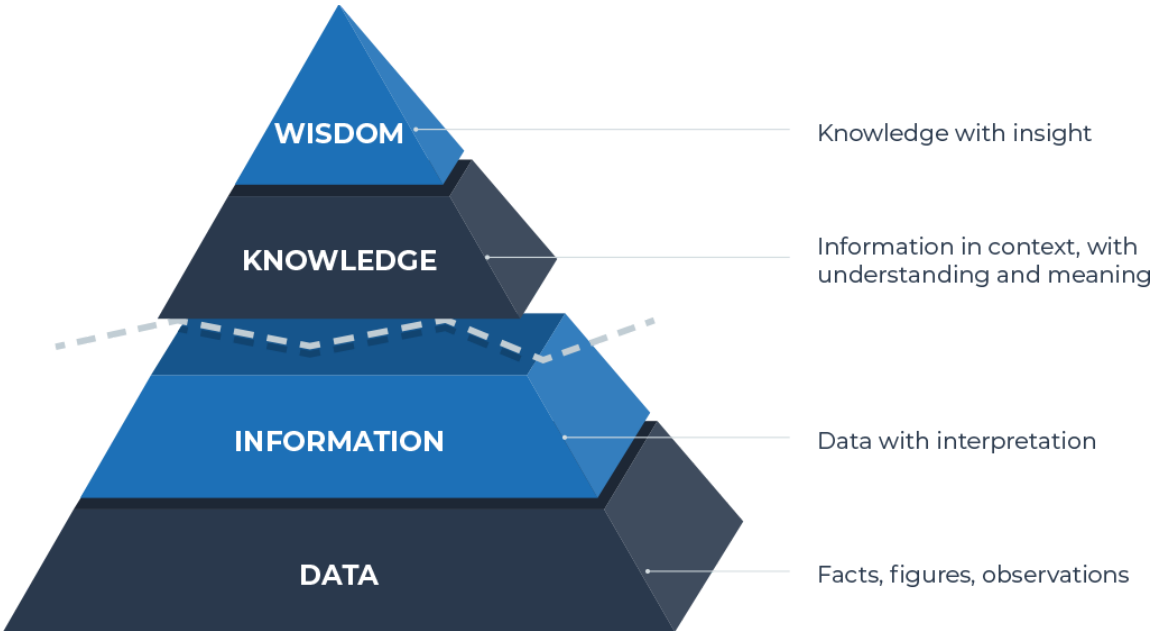
[5] T.H.Davenport & D.J. Patil, "Data Scientist: the Sexiest Job of the 21st Century", Harvard Business Review (October 2012) 73

[6] See Sloan Review for a condensed version of this paper - https://sloanreview.mit.edu/article/why-your-company-needs-data-translators/

[7] Op cit Davenport

In terms of the traditional knowledge hierarchy, the 'interpreter' function needs to operate primarily in the space between information and knowledge (understanding 'patterns') where information is transformed into knowledge (see Diagram). It is worth noting here that the terms 'data' and 'information' are being used almost synonymously in the debate at large and even confused with artificial intelligence. In sports and business, data is rarely collected without a reason and as such is immediately transformed into information. The 'translator' performs the *human* element in the hierarchy. As Zeleny explains,

> *while data and information can be generated per se, i.e. without direct human interpretation, knowledge, and wisdom cannot: they are human and context-dependent and cannot be contemplated without involving human comparison, decision making, and judgment* [8].



Where are we now?

Big data 'cheerleaders' argue that only numbers matter. Admittedly the availability of data has expanded exponentially; nevertheless the likelihood that such an explosion of data will fundamentally revolutionize analysis is under serious challenge. The more gung-ho claims such as those in Chris Anderson's 'Wired' article of 2008 where he claimed that, "the sheer volume of data will obviate the need for theory and even scientific method" and that, *"with enough data, the numbers speak for themselves"* are considered by such mathematical notables as Cambridge University's David Speigelhalter as, *"complete bollocks; absolute nonsense"*.[9]

The numbers may well speak for themselves, but they will always need translating into language that the decision-makers understand. Such translation may, indeed, help to avoid what is being referred to as data hubris - *"the often-implicit assumption that big data are a*

[8] M. Zeleny, 1987, "Management support systems: Towards integrated knowledge management", Human Systems Management 7 (1987): 59-70
[9] Quoted by T. Harford, "Big data: are we making a big mistake"? Financial Times, March 28, 2014

substitute for, rather than a supplement to, traditional data collection and analysis".[10] Essentially, this hubris is a result of the confusion between the power of correlation and the realities of causation. Tim Harford, the British economist, explains it simply, *"if you have no idea what is behind the correlation, you have no idea what might cause that correlation to break down."*[11]

With the sheer volume of available data comes an even greater increase in distinguishing between signals and noise. In his book on the same issue, Nate Silver, explains that,

> *If the quantity of information is increasing by nearly 3 quintillion bytes per day, the amount of useful information almost certainly isn't. Most of it is just noise, and the noise is increasing faster than the signals. There are so many hypotheses to test, so many datasets to mine - but a relatively constant amount of objective truth*[12].

It is clear, therefore, that the failure of big data to deliver is one of managerial and not technological inadequacy. The primary concerns for the sporting, corporate and public sectors are what an Accenture report describes as an analytics capability gap[13]. The report argues that *"practitioners are confident that analytics can deliver value but are frustrated that current capacity falls so far short of their aspirations." It also states there is "a real demand to get both more insight and more intelligence, but it's hard to find these people."* Another survey[14] confirms these worries about the scarcity of analytics talent, finding that, "55% of big data projects are abandoned"[15] and the top two reasons given by 80% of the respondents as to why analytics projects fail are that managers lack the right expertise in-house to "connect the dots" around data and, consequently, to form appropriate insights. There is an 'interpretation' gap, not a capability gap.

At the heart of the debate is a false dichotomy between the numbers and intuition. It is a dichotomy perhaps best understood through the counter messages of two movies, *Moneyball* and *Trouble with the Curve*, both of which examine the role of analytics in baseball. John Tumminia, a Chicago White Sox scout of over 30 years and possibly the inspiration for one of the movies, perfectly describes both the dichotomy and the solution. Which movie, we asked Tumminia, is closer to the truth about baseball scouts, *Moneyball* or *Trouble with the Curve*?

> *For me that movie [Trouble With The Curve] is very true because there are old time scouts who may or may not have been able to keep up with the new technology; they are set in their ways and may not be so in tune with the analytics stuff but on the scouting part of it, as you remember from the movie, the scout was right on when it came to that kid. He knew that he had to prove to the front office, because their numbers weren't showing it, that the kid couldn't handle the curve. He wanted to demonstrate that you shouldn't be married to the statistics. You have to understand that the scout will have a 'keen eye' to make the evaluation of the ability and the talent of the prospect. 'Moneyball' is a factual movie with a lot of fiction in it but 'Trouble with the Curve' is a fictional movie with more truth in it.*[16]

[10] D. Lazer et al, "The Parable of Google Flu: Traps in Big Data Analysis", Science, vol 343, (Mar 14, 2014)
[11] T. Harford, "Big data: are we making a big mistake"? Financial Times, March 28, 2014
[12] Nate Silver, "The Signal and the Noise: The Art and Science of Prediction" (London: Penguin, 2012)
[13] Accenture and MIT's Operations Research Center, "Analytics in Action: Breakthroughs and Barriers on the road to ROI" (2015 update).
[14] InfoChimps, "A survey of 300 IT professionals" (2014)
[15]IBM Institute for Business value, "Analytics: The new path to value: How the smartest organizations are embedding analytics to transform insights into action", Executive Report 2010.
[16] J. Tumminia, interview with authors, July 2014

Tumminia believes that balancing the old and the new is the answer to the best recruitment processes.

> *Obviously, the numbers do tell the story in many cases but there are always exceptions. The scout's evaluations will normally be accurate and the numbers will tend to confirm those evaluations and strengthen them – the numbers can step up the scout's grades....I've never forgotten the words of wisdom from Roland Hemond, my White Sox colleague who said, 'Learn how to change your report. Don't fall in love with a prospect so deeply that you cannot see his flaws'.*[17]

Tumminia's comments completely tally with those of R.C.Buford (San Antonio Spurs) who describes the decision making process as one that,

> *seeks alignment of the multi-variables – the eyes (scouting), the ears (intelligence), the numbers (performance). If we're going to deviate from what the eyes, ears, and numbers are telling us then there needs to be a really good reason. We would need to dig much deeper.*[18]

The sentiments expressed above signal the way ahead as being a symbiotic learning relationship between the numbers and intuition and that requires relationships between different personality types.

Why isn't it better?

Given that there is fairly universal agreement that big data can add value, why is it not doing so more consistently; why does it underperform? We do not believe that it is a 'capability gap' in the terms of the technological skills required, as many argue, rather it is a cultural gap in the manner described above by Tumminia and Buford which is simultaneously definitional, attitudinal, linguistic and structural.

Definitional

The nomenclature around this issue is gradually converging but it still needs some common-sense agreement. Before that can happen, the actual role itself needs to be agreed. Are we comfortable, for example, with the term 'data scientists'? Definitely not, if it is synonymous with the grandiose descriptions currently attached to the term. Even the use of 'scientist' can suggest distance from the core business. What about 'data analyst'? Probably more accurate but the term 'data' still tends to be associated with the scientific in a potentially 'geeky' way. The easiest title might be 'intelligence analyst' which encapsulates all information not just the digital.

Attitudinal

The most obvious and possibly most pervasive attitudinal bias is what appears to be a natural human predisposition to resist change. It is the essence of the old school/new school confrontations that are evident in every sector. Other well researched biases also dampen the utility of data-driven intelligence such as the *overconfidence bias,* an individual's confidence in their own judgement and their own abilities which is regularly at odds with reality. However, in their article on the NFL draft, Massey and Thaler point out that,

[17] J. Tumminia, interview with authors, July 2014
[18] J. Tumminia, interview with authors, July 2014

> *human nature conspires to make it extremely difficult to avoid overconfidence in [draft selection]. The more information teams acquire about players, the more overconfident they will feel about their ability to make fine distinctions. Even the smartest guys in the world, the guys who spend hours with game film, can't predict [draft choices] with much success. There is no crime in that. The crime in thinking you can predict it.* [19]

Cognitive biases abound and Billy Beane describes the *emotional bias* that sports people are under as being of major significance. He believes that,

> *All decisions are now public decisions, everyone is an expert. In sport it's continuous at least in business it is usually only quarterly. There is permanent media scrutiny and it must have some sort of effect and decision making. The decision maker needs to eliminate the noise because it creates an emotional bias.*[20]

Despite the numerous biases, things do change, there is progress. Attitudes can change but it's an uphill battle against the myriad biases that we all carry. As Prasad Setty, the VP People Analytics, at Google, states, *"it's not about replacing humans but about reducing bias*".[21]

Linguistic

An equally powerful theme from our research is the significance of the communication barrier. It is apparent that senior management, for whatever reason, in whatever sector, do not speak the same language as the analysts. Decision-makers are said to be seeking better ways to receive complex insights in a language that enables them to easily absorb the meaning of the data. An IBM survey suggests that,

> *Over the next two years, executives say they will focus on supplementing standard historical reporting with emerging approaches that make information come alive. These include data visualization and process simulation, as well as text and voice analytics, social media analysis, and other predictive and prescriptive techniques.*[22]

In simple terms that means, in a form they can understand.

A blog written by Chad Thompson, a self- confessed ' geek', hits the nail on the head. He explains that the import of a Gartner report on the barriers to big data utility shows that,

> *the biggest issues in technology have little to do with technology, and more to do with defining the meaning of language. It's impossible to make sense of data unless there is some baseline agreement on what the data should actually* ***mean****. It should be obvious that words can have radically different meaning from one group to another – it takes the skill of collaborative effort to* ***define what words mean for an organization****.....Dealing with business data can be even more difficult because there may be ingrained cultural differences in organizations – the "IT Department", the "Sales Department", the "Marketing Department", etc..*[23]

Sig Mejdal, Vice President and Assistant General Manager, Baltimore Orioles argues that most decision-makers are, "not conversant with the scientific method, so we have to change

---

[19] C. Massey and R. H. Thaler, "The Loser's Curse - Decision Making and Market Efficiency in the NFL Draft", Management Science, Articles in Advance (2013 INFORMS): 1–17,
[20] B. Beane, interview with authors, January 2015.
[21] MIT Sloan Sports Analytics conference, Boston, February 2015
[22] IBM Institute for Business value, "Analytics: The new path to value: How the smartest organizations are embedding analytics to transform insights into action", Executive Report 2010.
[23] http://chadthompson.me/2013/01/; 15th January 2013

our language". People who can do that are hard to find because those educated in the scientific method find it difficult to alter their language to make it more accessible. Similarly, senior management finds it difficult to explain to the scientists that often 'enough is good'; that 'answers often can't wait for perfection'. Functionally, some form of translation is needed, and it can't be computer-based; it needs a human translator.

Structural

Structurally, this function has to bridge the intellectual space between the information and knowledge layers of the accepted knowledge hierarchy. Formally, this could be housed in a business intelligence unit but in reality, the formality is not the important part. It is the type of person that inhabits the function that is key to its success or otherwise.

Who does this role well; who gets it?

When the symbiotic relationship between the data and the management is done well it looks easy. However, the first part of the process is to understand both sides of the equation. What those who currently do this well are able to do, is simultaneously be the management (the domain specialist) and the bridge (the translator). The 2016 MIT sports analytics conference had a powerful panel of those who 'get it'. Del Harris, a legendary NBA head coach with the Houston Rockets, Milwaukee Bucks, and Los Angeles Lakers, was, in an earlier incarnation, the translator himself. He explained the role precisely – he said, "during my time with the Mavericks the analytics used to come directly to me and the owner [Mark Cuban] because Don Nelson [the head coach] couldn't care less about that sort of thing". Harris would then 'translate' the data into questions and insights using language with which Nelson was comfortable.

Eric Mangini, the former Head Coach of the Cleveland Browns and New York Jets was asked how analysts can grab the attention of the decision-makers. His simple answer, "show your results and show how they can help us win games". RC Buford is equally precise. "When the alignment between the eyes, ears, and numbers isn't there, falling back on the process and the data allows you to ask better questions".[24] The translator function demands that the analysts generate radical questions because the domain specialists will almost inevitably know the answers. Also, inevitably, they will have mostly failed to ask themselves the questions. Conversely, McAfee & Brynjolfsson argue that,

> *When it comes to knowing which problems to tackle, of course, domain expertise remains critical. Traditional domain experts - those deeply familiar with an area – are the ones who know where the biggest opportunities and challenges lie.*[25]

We believe that they are so deeply involved in the game/business operations that they fail to ask the outlying questions most likely to unlock innovatory thinking.

Others who 'get it' consistently recite the mantra that the key to driving analytical change is to be able to trust in the support from the boardroom of your organization; to use data and analytics to provide fact-based evidence for managerial approaches and decisions; to have the determination during the times when results are affected by external/mitigating circumstances to stay the course.[26]

---

[24] R.C. Buford, interview with authors, December 2014

[25] A. McAfee & E. Brynjolfsson, "Big Data: the management Revolution", Harvard Business Review, special edition (October 2012): 66.

[26] Some useful discussion around the issue of trust is contained in the following article - Sarah Bayer, Henner Gimpel & Moritz Markgraf (2021) The role of domain expertise in trusting and following explainable AI decision support systems, Journal of Decision systems, DOI: 10.1080/12460125.2021.1958505

Billy Beane echoes these sentiments.

> *It's absolutely central to have a manager who buys into what the front office is doing. A manager who is open to new ideas; hiring Scouts and number crunchers who can work together instead of against each other; and recruiting 'quants' who know how to explain data in plain English so that the manager and the other field personnel can easily put that information into play.*[27]

Beane, unequivocally, pointed to the Oakland manager Bob Melvin as being virtually unique in bridging that gap. He believes that Melvin's specific background as a professional player (a domain specialist) and a Wall Street quant (an analyst) enables him to perform a unique function with the As. He is the bridge, and it makes the point precisely that it is a domain specialist who 'gets' the quants who can best fulfil the translator function. Interestingly, two very successful head coaches, Eric Spoelstra of the Miami Heat and Jose Mourinho of AS Roma are also individuals who have knowledge of all aspects of the organisation. Indeed, Mourinho was an 'actual' language translator in his early days and Spoelstra started at the Heat as a video coordinator, then assistant coach, then advance coach, then finally succeeding Pat Riley as head coach.

Finally, from a sport-related corporate organisation, the co-founder of Bleacher Report, Dave Finocchio, explains how important the 'translator' layer is.

> *It's important to have a translator layer. The domain specialists are more valuable and have more insight than the trained data scientists. Content is very personalised, you have to have innate knowledge to be an expert in the space. So, what we've done at Bleacher Report is to take people with an aptitude for understanding data from the content team and we've made them members of the analytics team.....We made our analytics team God, everything stems from them but they have to be strong from both a data standpoint and content and that's why we migrate them from the domain to the analytics team. To say that the analytics team effectively generate the questions definitely resonates with me. The domain specialists have to really understand the data and they have to know that the data they are looking at is accurate. It's part art and part science."*[28]

What's the Solution?

There is clearly a need for the ideas, concepts and questions constantly generated by organisations' analytics teams to be translated into formats, be they written, aural or visual, with which the decision-makers are able to engage. It is important to understand that this is a function and not a single person. This clearly isn't happening in enough organisations to make a significant impact. This may be because organisations are still looking for the mythically talented data scientist described elsewhere. Greta Roberts, CEO and co-founder of Talent Analytics Corp. says that the biggest problem is that the current job description doesn't hit the mark.

> *It's over-specified. There is a null set of people that fit the entire description. They're unicorns; you can't find them. Or there are a very limited number of people that fit the criteria. When you review data scientist hiring criteria, you'll find mutually exclusive requirements. They want charismatic communicators that are able to effectively present findings. At the same time, they want people to sit and work with*

[27] B. Beane, interview with author
[28] D. Finocchio, interview with author, 7 October 2014, London

data all day. These are two different types of people. Our data shows companies in fact split up these roles.[29]

In fact, splitting these roles and the consequent silo-ism is a serious problem. Identifying the necessity of the interpretation function is actually the most crucial step but the second step is to develop a personality characteristics map for the function. As Butch Cassidy once said to Sundance, "Who are these guys?" The personal attributes required for 'translators' to bridge the interpretation gap are the ones that need to be identified in the recruitment process. The translators as we envisage them require simple but elusive skills:

- sufficient domain knowledge to pass the 'street cred' test with the head coach/CEO/COO/decision-maker.
- an ability and willingness to learn sufficient analytics to be able to talk the talk.
- the confidence to speak truth to power and to peers and to subordinates.
- to have a cognitive "attitude" and a willingness to search for deeper knowledge about everything.
- a creative drive and need to create not only questions but also potential answers in a form that others find accessible.
- a willingness to embrace the organizational culture regarding work ethic.
- an extremely high sense of quality, standards, and detail orientation.
- the ability to robustly engage at team or organizational meetings without being asked for input and even if the topic is one of high importance.

Although the personality traits will tend to be innate, the process that the translators must follow to deliver results can be taught. For example, they need to,

- *Connect with the reactionaries through questions, not assertions.* In an alien environment, it is essential not to be overly assertive at the outset. Ask questions that enable the decision-makers to come up with the answer, ostensibly by themselves.
- Create analogies around anecdotes that resonate with the decision-makers. These are the stories of successful analytic interventions such as, *'this guy wasn't expected to make it and she has' or 'this was counter-intuitive, but it worked'.* It is important to get the decision-makers to actually experience the anecdotes.

[29]Predictive Analytics World Conference, "Benchmarking Analytical Talent": April 2013

Conclusion

There is an obvious need for a symbiotic relationship between the domain specialists and the analytics specialists. As even the high priest of analytics, Billy Beane, acknowledges, *"I used to think at one stage that an actuary could do my job but I realize that you can't simply ignore the emotional side in such an emotional business."* It is only partially true that you, *'can't manage what you don't measure'* but it is completely true that, *'not everything that can be counted counts and not everything that counts can be counted'*. Bridging the cultural gap with an 'interpreter' function performed by a 'translator' can begin to address the disparity between the claims for big data and its reality.

In turn, that process begins with recognizing the limitations of what numbers and intuition can do separately. While that recognition is grudgingly emerging the current expectation that 'data scientists' can provide the solution is misplaced. Even though we actually know that C-3PO doesn't really exist we can at least try to emulate his primary function as a translator - *to enable ventures between different cultures to run smoothly*. And, there aren't many working cultures more different than those of the quants and their managers. To give the last word to Billy Beane,

> *It's just data - it's how you interpret it and then what you do with it that matters.*

For more information on this topic, or to engage directly with our research team, contact:

Professor Chris Brady
Chief Intelligence Officer - Sportsology
chris.brady@sporstology.pro
+44 (0)7956575278