Impact of Big Data Analytics on Decision Making and Performance

Authors: *Usarat Thirathon** *Usarat.Thirathon@uts.edu.au*
Bernhard Wieder *Bernhard.Wieder@uts.edu.au*
Zoltan Matolcsy *Zoltan.Matolcsy@uts.edu.au*
Maria-Luise Ossimitz *Maria.Ossimitz@uts.edu.au*

Affiliation: *University of Technology Sydney, Accounting Discipline Group*

Abstract:

'Big Data' has become a major topic of interest and discussion for both academics and professionals in the IT and business disciplines, and case evidence suggests that companies engaging in Big Data outperform others. It has to be noted though that 'Bigger' Data as such does not provide any benefits, but it is rather how organisations make sense of data and gain insights from analysing the data. Analytic capabilities and practices are required to convert Big Data (BD) into insights which arguably improve decision-making and thereby organisational performance. While protagonists of such Big Data Analytics (BDA) imply that those effects exist, so far they have not been confirmed by rigorous empirical research.

Data was obtained using a cross-sectional online survey which targeted Chief Information Officers and senior IT managers of medium-to-large Australian for-profit organisations and yielded 163 complete responses, which met the standard criteria for measurement reliability and validity. PLS-SEM and multiple bootstrapping methods were used to test the hypotheses, while controlling for firm size.

The present study empirically confirms claims made in the literature that BD and related analytics lead to better performance. It also reveals that such benefits are achieved primarily because BDA creates additional incentives for managers to base their decisions on analytics, and that more *analytic-based* decision making actually leads to superior performance. Finally, the results of our study suggest that managers in organisations which engage in BD are generally more analytics-minded in their decision making, even if the analytic tools and methods used in support of their decisions are not particularly sophisticated.

The results provide evidence that neither Big Data nor Big Data Analytics are just 'hypes', but they do actually lead to superior performance, partly directly and partly indirectly by creating an incentive for managers to rely on analytics when making strategic or operational decisions. Interestingly, managers in smaller firms are more likely to base their decisions on analytics than larger firms, which suggests that they use analytics to compete against larger firms.

Keywords: Big Data, Big Data Analytics, Organisational Performance, Decision Making, Benefits of Analytics

* *Corresponding author*
[1] *UTS Business School, University of Technology Sydney, Ultimo, NSW, Australia*

1 Introduction

Over the past decade, 'Big Data' (BD) has become a very popular term in practitioner and academic conferences, journals and books. In essence, the term roots in the fact that data generated and available today is 'big' in terms of volume, variety, and velocity (Chen et al. 2012, Davenport 2014). The broader meaning of the term also includes all the analysis performed on the data (Big Data Analytics).[1] Executives increasingly acknowledge the potential benefits associated with BD (Schroeck et al. 2012) and high-performing organizations believe that BD is a critical differentiator and a key to growth (LaValle et al. 2011, Schroeck et al. 2012). It has been claimed that BD is a paradigm shift that not only changes the way organisations deal with data, but also the way they run their businesses (Vasarhelyi et al. 2015). Based on those promises, global private and public investment in BD has reached billions of dollars per annum (Gartner Research 2015, Rossino 2015).

It is obvious that generating and storing vast amounts of diverse and volatile data comes at a cost, but per se does not provide any benefits. It is rather the *insights* and *decision support* gained from analysing the data which allows organisations to compete on analytics (Davenport 2006, King 2013). The claimed power of BD does not replace the need for human insight (McAfee and Brynjolfsson 2012), but supported by BDA experts (data scientists), who can provide such insights from data, managers are expected to make better (informed) decisions than without BD and BDA (Davenport and Patil 2012, Bange and Janoschek 2014, Davenport 2014). On the other hand, executives still struggle to understand and implement BD strategies effectively (CGMA 2013), and some argue that the biggest challenge with BD/BDA is that managers do not comprehend how to gain benefits from analytics (LaValle et al. 2011).

So while there are convincing BD/BDA success stories, there are substantial costs and other challenges involved with BD (initiative), and so far there has been no larger scale empirical evidence of net benefits. Accordingly, our first research question is: *Does Big Data Analytics lead to superior performance and therefore competitive advantage?*

The second research question explores the mechanisms ('how') through which organisations may create competitive advantage with BD/BDA. BDA is about providing new insights (*knowledge discovery*) and *decision support*, i.e. the business case for BDA investments is effectively about better and faster decisions (including decisions about business models and processes), which have proven to lead to better performance (Meissner and Wulf 2014, Wieder and Ossimitz 2015). But all this assumes – amongst others – that managers actually *use* available BDA output to support their decisions. As mentioned before, this requires that managers actually understand the implications of the analytic output (LaValle et al. 2011), and have the ability and willingness to uncover and use insights from BDA (Deloitte 2014). Behavioural decision theory offers a range of explanations why managers may ignore analytic evidence or advice (Edwards 1961, March 1978) and use other decision criteria (incl. intuition or simply 'politics'). Accordingly, the second research question in this paper is: *Are managers in organizations with sophisticated Big Data Analytics more likely to base their decisions on analytics (facts, evidence) than managers in organizations low on Big Data Analytics, and if so, does this lead to superior performance?*

[1] For the purpose of this research, we distinguish clearly between the data dimension (BD) and the analytic dimension (BD Analytics - BDA).

Finally, we explore the relationship between BD and BDA, and how this relationship may influence the answers to the above-mentioned research questions. We have already established that BD is useless without analytics, so it is reasonable to expect that organisations with high BD 'intensity' will have more sophisticated BDA; but we also have to ask the reverse question: *To what extent is sophisticated BDA useful or beneficial for relatively 'small' data?* It is e.g. argued that real-time text mining can discover actionable and meaningful patterns, profiles, and trends from text/web resources (Linoff and Berry 2001) and address data that is streaming continuously on social media (Chakraborty et al. 2013). But one might also argue that organisations which mine only structured data achieve equally good outcomes.

The objectives of our research are as follows: First, we want to empirically verify claims made primarily in the practitioner literature that BD and its analytics (BDA) leads to better performance and competitive advantage. Second, our research intends to verify to what extent any such benefits are achieved because BDA creates additional incentives for managers to base their decisions more on analytics (analytic-based decision making – ABDM)[2]. Finally, we intend to uncover any mediating and/or moderating effects of BD intensity on the above relationships.

The reminder of the paper is organized as follows: Section 2 elaborates on the constructs of interest and makes predictions about their relationships (hypotheses); section 3 explains the research method, including construct measurement, and section 4 presents the results. Finally, the results, their implications and the limitations of our research are discussed in section 5.

2 Theory/Hypotheses Development

2.1 Background

Big Data (BD) refers to a set of techniques and technologies that require new forms of integration in order to uncover hidden value from large datasets that are diverse, complex, and of a very large scale. The *volume* of BD is massive, so conventional hardware and software are incapable of handling it within a suitable time-frame (Moffitt and Vasarhelyi 2013). Data volume has increased dramatically and the unit of measuring data will change from zettabytes (1021 bytes) in 2012 (Davenport 2014) to yottabytes (1024 bytes) in 2030 (IEEE 2013). The *variety* of data used for analytics also increased dramatically, because it includes not only traditional relational data, but also raw, semi-structured, and unstructured data from various sources. Unstructured data, such as emails, text-based documents, images, videos, call-centre recordings, and sensor-generated data cannot be stored easily within a standard relational database (CGMA 2013) and require new analysis techniques. *Velocity* refers to the speed of both data generation and data processing. Data generation is rapidly increasing as a result of widely-used mobile technologies and 'The Internet of Things'. Real-time or near real-time information are said to enable organizations to be more agile than their competitors (McAfee and Brynjolfsson 2012). Data today is generated, changed, and removed substantially more frequently than in the past and consequently organizations need new platforms and tools for analysing them.

[2] In practice, the term 'data-driven' decision making is much more common, but considering we clearly distinguish data from data analytics, analytic-based decision making is clearly a more appropriate term.

"Analytics is the science of analysis" (Turban et al. 2008). Data analytics[3] uses data for quantitative and/or qualitative analysis to help an organization better understand its business and markets (knowledge discovery) and to make timely business decisions (Chen et al. 2012, Holsapple et al. 2014). Data analytics involves multiple disciplines, in particular, mathematics and statistics, but also data mining, business intelligence (BI), machine learning, pattern recognition, and data visualization. Analytics can be descriptive, predictive, and prescriptive in nature (Chen et al. 2012, Minelli et al. 2012, Davenport 2014). *Descriptive analytics* drills down into past or current data to discover trends or patterns to support managerial decisions. *Predictive analytics* supports organisational decisions and strategies by gathering historical data, forecasting, and simulating to anticipate possible future situations. *Prescriptive analytics* refers to descriptive and predictive analysis of data that suggests a set of potential actions to managers considering rules, constraints, thresholds, risks, and uncertainty. Prescriptive analytics provides the most concrete decision support, and considering that the trend in analytics is moving from historical analysis to forward-looking predictive and prescriptive analytics (Hagel 2015), one would expect increasing *decision-relevance* of analytics.

Data analytics in a BD environment, i.e. Big Data Analytics (BDA), is different from conventional data analytics for the reasons mentioned above (McAfee and Brynjolfsson 2012). With the emergence of BD, the analytic algorithms have changed so as to be able to deal with the high volume, variety, and velocity of data. BDA applies new scientific methods to solve problems that were previously impossible to solve, because either the data or the analytic tools did not exist (Davenport 2014, Parmar et al. 2014). BDA no longer involves just traditional hypotheses-based statistical analysis, but also machine learning, predictive modelling, faster processing tools, high-performance analytics environments, and visual analytics (Chen et al. 2012, Dhar 2013, Dyché 2014).

2.2 Hypotheses

Descriptive and predictive analytics primarily contribute to knowledge discovery thereby adding to the general pool of knowledge available in an organisation, which managers can derive decisions from. Prescriptive analytics, however, provides direct decision support or can even make automatic decisions without managerial intervention. However, analytics cannot guarantee that decisions made are 'optimal' for many reasons, including:

a) Analytic insights or recommendations are ignored or not prioritised in decision making;
b) Analytic outcomes are misinterpreted in decision making (especially outcomes of descriptive and predictive analytics);
c) Analytic models are never a perfect representation of reality,[4] let alone the future;
d) The objectives in decision problems are unclear or conflicting;
e) Analytic algorithms are used improperly; etc.

The very basic condition for any potential positive effects of analytics are mentioned in limitation a) above: Managers have to actually 'listen to' analytics, i.e. incorporate descriptive and predictive insights into their decisions, act upon prescriptive analytics or let the latter decide automatically.

---

[3] Data analytics is a sub-field of the broader concept of *data science*, which – in the broadest sense – "develops relevant methodologies, theories, technologies and applications for data, ranging from data capture, creation, representation, storage, search, sharing, privacy, security, modeling, analysis, learning, presentation and visualization, to integration across heterogeneous, interdependent complex resources for real-time decision-making, collaboration, value creation, and decision-support" (Cao 2016).

[4] Constructivist epistemology negates any 'objective reality' (Watzlawick 1984).

Behavioural decision theory offers a range of explanations why managers may ignore analytic evidence or advice (Edwards 1961, March 1978). Davis (1989), on the other hand, offers a theory which explains the acceptance of information technology in relation to the perceived usefulness and ease of use of the technology. While analytic output is not technology, it is based on information (technology) processes, and if it is useful in terms of overcoming bounded rationality (Simon 1972) or has proven useful in the past, and is further properly presented and applicable (= easy to use), managers will be more likely to use analytic evidence or advice.

But even if managers base their decisions on analytics, there are still many other reasons why decisions are 'sub-optimal' (e.g. limitations b-e). Despite all these limitations, analytics can help to make better *informed* decisions, which should lead to better decisions and thereby improved performance.[5] BDA tools and the methods embedded can therefore help organizations create actionable strategies by providing predictive and prescriptive analytics, which provide deeper insights in how to address their business requirements and plans (Barské-Erdogan 2014). With new technologies and analytic approaches, BDA can provide managers with information for real-time planning and continuous forecasting (McAfee and Brynjolfsson 2012, Moffitt and Vasarhelyi 2013, Barské-Erdogan 2014). BDA techniques are capable of analysing larger amounts and types of data with increasingly advanced algorithms, which allow more prescriptive analytics. With such 'easy to use' information, managers are expected to act more on analytics and improve decision efficiency and effectiveness (Brown-Liburd et al. 2015).

In summary, we predict that BDA sophistication leads to superior performance (a) indirectly, by producing more timely, relevant and actionable information thereby creating an incentive for managers to act upon that information for superior performance and – to a lesser extent – (b) directly, by means of automation of decisions and business processes.

H1: Big Data Analytics sophistication has a positive total effect on organisational performance.

H2: The positive effect of Big Data Analytics sophistication on organisational performance is partly mediated by analytic-based decision-making.

Some organizational resources are *complementary,* i.e. they have to work together with other resources to maximize their impact (Barney 1986). The relationship between BD intensity and BDA in affecting ABDM is proposed as an example of such complementarities (Masli et al. 2011). BD intensity in itself is not expected to have a positive effect on ABDM. Just because an organisation has access to/uses a greater volume and variety of rapidly changing data will not per se create an incentive for managers to use that data for decision making – more likely the reverse: In the absence of proper BDA, BD intensity will more likely result in information overload, which will make it more difficult to understand and process the data (Yang et al. 2003, Pfeffer and Sutton 2006, Rousseau 2006), and more likely deter managers from using it. However, when more sophisticated analytic tools and methods are used with 'bigger' data, the analytic outputs are expected to offer greater insights than BDA based on smaller and less diverse data sets, and will therefore be used more for decision making. So while BD intensity is expected to have a weak – if not negative – direct impact on ABDM, its impact on ABDM in combination with BDA is expected to be positive.

[5] Of course, any such 'good' decisions still need to be implemented to become effective; we considered this fact in the design of our research instrument, in which we included a question about 'acting on insights' for measuring ABDM.

This predicted complementary effect can be interpreted and hypothesised in three ways: (a) as a mediating (indirect) effect of BD on ABDM via BDA, (b) as interaction effect between BD and BDA on ABDM, and (c) as moderated-mediation effect as a combination of (a) and (b).

a) The availability of greater volumes of more diverse and volatile data creates an incentive for deploying more advanced analytics to 'make sense' to the data, e.g. through advanced visualisation. Considering that BD relies on BDA to be potentially useful, we predict:

H3a: Higher Big Data intensity leads to more sophisticated Big Data Analytics and – subject to confirmation of H1 and H2 - has a positive indirect effect on analytic-based decision-making and performance.

b) Alternatively, it could be argued that sophisticated BDA is of little value for decision makers in organisations which are low on BD, whereas organisations which *do* engage in BD are expected to benefit substantially from using advanced BDA tools and methods to gain insights from BD (see Figure 1, chart 1); or in other words: The higher the BD intensity, the stronger the impact of BDA on ABDM. In this alternative explanation, BD intensity is expected to have a *positive moderating* effect on the relationship between BDA and ABDM.

H3b: The positive relationship between Big Data Analytics sophistication and analytic-based decision making is positively moderated by Big Data intensity.

c) Finally, it can be proposed that the moderation effect as predicted in H3b does not just interact with the relationship between BDA and ABDM, but also the whole indirect effect path as predicted in H3a; such a scenario is referred to as moderated mediation (Preacher et al. 2007, Hayes 2015).

H3c: The indirect effect between Big Data intensity and analytic-based decision making (via Big Data Analytics) is moderated by Big Data intensity.

Figure 1 – Big Data and Analytics Scenarios

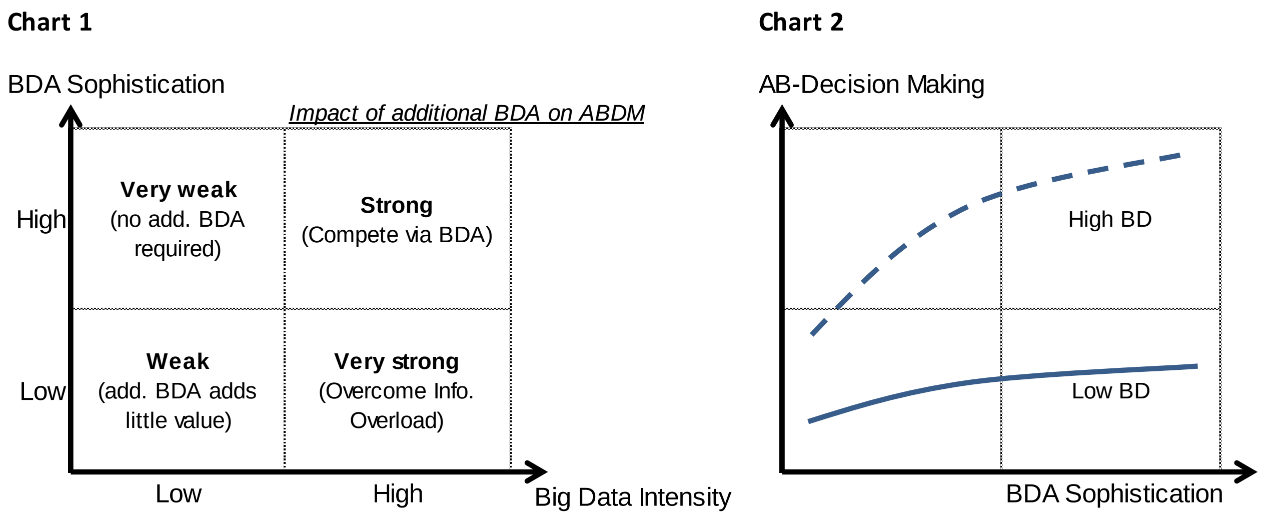

Chart 1 depicts four BD-BDA scenarios and describes the marginal impact of BDA on ABDM for each of the scenarios. Chart 2 'translates' the scenarios into stylized regression curves; in line with the effects predicted in chart 1, the regressions are non-linear.

3 Research method

As this research is exploratory in nature, a cross-sectional survey has been selected as most suitable research method (Rindfleisch et al. 2008). Survey procedures used follow the suggestions of Dillman et al. (2014). As each variable in the hypotheses is latent, constructing proper indicators and scales was essential. This process was guided by academic literature, but where required, practitioner literature was also consulted. During questionnaire design, necessary procedural remedies were applied to control for and minimize the impact from these common method biases (Podsakoff et al. 2003).

The *face* and *content validity* of the prototype of the questionnaire as well as the appropriateness of Likert scale endpoints were assessed as follows (Podsakoff et al. 2003): Five experts in survey research were invited to evaluate the draft questionnaire, and their feedback was used to refine the design and content of the survey. The revised version of the questionnaire was then delivered to a small sample group of experts for pilot testing. The survey targeted CIOs and senior IT managers of Australia-based medium to large for-profit organizations.

3.1 Measurement

In this research, *BDA sophistication* is operationalized along two dimensions (Acito and Khatri 2014): (a) analytic tools and (b) analytic methods. *Analytic tools* refer to software applications that analytic professionals use in data analytics. They range from basic spreadsheets to business intelligence (BI) tools, statistical packages, data mining suites, data visualization tools, and high performance computing tools. *Analytic methods* refer to quantitative methods that analytic professionals use in data analytics. They include statistical methods, machine learning, data mining, artificial intelligence, operations research, optimization models, and path modelling (Dhar 2013). Respondents were asked to rate their analytics expert/team in terms of various analytic tools and methods (skills) on a seven-point Likert scale in terms of frequency of use of each analytic tool or method, with 1 = never and 7 = very frequently (see Table 2 in the appendix).

To measure BD a three-dimensional scale for 'BD intensity' was developed based on the three Vs (volume, variety, and velocity). Respondents were asked to rate the level of *increase* of each 'V' in terms of data (a) they have *access to* and (b) data they actually *use* in analytics in their organization using a five-point Likert scale (see Table 2 in the appendix).

In an organization, *decisions* can be made at many levels: strategic, tactical, and operational (Nowduri 2011). Strategic decisions include how an organization initiates new products, services, or market channels, which major suppliers it selects, etc. Operational decisions involve day-to-day decisions in various business functions, e.g. marketing, operations, and procurement.[6] Respondents were asked to rate the level of their organizational strategic and operational decisions relying on insights derived from data analysis/analytics on a seven-point Likert scale, in which 1 = strongly disagree and 7 = strongly agree.

Organisational performance is measured relative to the performance of the major competitor (Garg et al. 2003). When using archival data, researchers usually measure performance either with market performance measures (stock market return, Tobin's q) or accounting performance measures (profit

[6] Tactical decisions were excluded from the questionnaire, because they were deemed to be in a too 'grey area' between strategic and operational decisions.

margin, turnover ratios) (Dehning and Richardson 2002). Performance can be measured either at the business process level (operational efficiency) or the organisation level (productivity, efficiency, profitability, market value) (Melville et al. 2004). Four performance indicators were used: sales growth, market share, profitability (Peters et al. 2016), and cost reduction. Respondents were asked to rate their organisation's performance relative to their main competitor, in the past 12 months (seven-point Likert scale, in which 1 = much worse and 7 = much better).

Firm size was used as a control because it can affect decision-making rationality (Mintzberg and Waters 1982, Garg et al. 2003) and systematically influence organisational practices and organisational performance (Baum and Wally 2003, Garg et al. 2003). Firm size is measured using the number of full-time equivalent (FTE) employees.

3.2 Survey response

The initial invitation was sent out to 1,595 potential respondents via email, but 263 invitations did not reach the addressees. A total of 174 responses were received during the survey period, but 11 had to be excluded, because they did not meet the required criteria (in terms of minimum tenure etc.). The final response rate of 12.24% may appear low, but is not unusual in Australian business surveys, even more so on a topic which is still only emerging. The spread of the responses reflects Australian industry, as shown by the coefficient of variation (CoV). The CoV of the survey responses is .844 while that of Australian industry is .843. 84% of the responses came from organizations which have more than 100 full-time equivalent employees and CIOs and other senior IT managers were almost equally represented (52.1% and 47.9%).

3.3 Data characteristics and quality

Tests for normality were conducted for both indicator data and latent constructs in order to determine the appropriate analysis and testing techniques (parametric vs. non-parametric) (Kraska-Miller 2014). The *Shapiro-Wilk test* and the *Kolmogorov-Smirnov test* both show that none of the indicators is normally distributed ($p < 0.05$), which requires the use of non-parametric data analysis and testing techniques (PLS-SEM, bootstrapping, etc.) (Hair Jr et al. 2014).

In addition to the procedural remedies applied during the development of the questionnaire, post-hoc statistical remedies were used to test for potential *method bias* (Podsakoff et al. 2003). *Harman's single factor test* was run across the set of 32 measurement indicators and the results show that there are 7 factors with Eigenvalues greater than 1 indicating that common method variance due to method bias is not present.

Responses were also tested for *non-response bias* by comparing early and late responses. The results of independent samples test (*Mann-Whitney U* and *Levene's Test*) confirm that there are no significant differences in the indicator values between the early (n = 83) and late (n = 80) response group.

After the elimination of three low-loading indicators, all remaining indicators have significance levels of $p < .001$, and load primarily on their assigned construct (Table 8). The measurement model was further assessed for reliability and validity of the construct measures. Reflective measurement models are assessed for: (a) the internal consistency (composite reliability), (b) indicator reliability (composite reliability), (c) convergent validity (average variance extracted and communality), and (d) discriminant validity (Hair Jr et al. 2014).

Table 4 confirms that the first three of these criteria are fully met. The Fornell-Larcker criterion (Fornell and Larcker 1981) was applied to assess for *discriminant validity* of latent constructs both at the first (Table 6) and second (Table 7) order level, and all constructs meet the criterion. The heterotrait-monotrait (HTMT) ratio between the average of the heterotrait-heteromethod correlations and the average of the monotrait-heteromethod correlations (Henseler et al. 2015) is used to further ensure discriminant validity. A HTMT value of two latent constructs of less than .85 confirms discriminant validity between the pair. Table 9 reveals that there are some measurement 'overlaps' between the '3 Vs' of BD, but considering that they are used to form a second order formative construct with satisfactory VIF-scores (Table 3), this minor lack of discriminant validity is not considered a concern.

4 Results

Two slightly different versions of the structural model were developed to test the hypotheses. The first model (A) was used to test hypotheses 1, 2 and 3a. Model B was used to determine the moderating effects as predicted in hypotheses 3b and 3c, i.e. BD and the interaction term BD*BDA were added to the model. The results of the *PLS analysis* and *bias-corrected bootstrapping* are presented in Table 1.[7]

Model A: As predicted in hypothesis 1, BDA sophistication has a significant positive total effect on organisational performance ($\beta = .254$, $p < .001$). The test results also confirm hypothesis 2, because the indirect effect of BDA on performance via ABDM is also significant ($\beta = .116$, $p < .01$). The strong indirect effect is the result of strong direct effects between BDA and ABDM ($\beta = .305$, $p < .001$, f = .105*) and between ABDM and performance ($\beta = .380$, $p < .001$, f = .146*). Although not hypothesised, the results also reveal a significant direct effect of BDA on organisational performance ($\beta = .138$, $p < .05$, f = .019), although the size of the effect (*f* square) is too small to be significant. While the *beta* of that indirect effect ($\beta = .116$) is marginally smaller than that of the direct effect ($\beta = .138$), the latter is more significant and is characterised by much higher effect sizes (*f*) along the indirect path.

Model A also confirms hypotheses 3a by revealing a significant indirect effect of BD on both ABDM ($\beta = .123$, $p < .001$) and performance ($\beta = .210$, $p < .001$). The latter is stronger than the former as it also includes the direct impact of BDA on performance. However, model A also reveals a very strong direct effect of BD on ABDM – very much in contrast to our expectations that any such effect would be fully mediated by the indirect effect via BDA. This surprising result has implications for the results for model B.

Model B extends model A by introducing an interaction effect between BD and BDA (on ABDM), but in contrast to our predication, such effect is not significant (and even has a negative sign), thereby rejecting hypothesis 3b – and thereby also hypothesis 3c. In fact, the inclusion of the interaction term changes hardly any result compared to the values in model A. The absence of interaction between BD and BDA is, however, consistent with the unexpected strong direct effect of BD on ABDM already apparent in model A. It appears that high levels of both BD and BDA drive ABDM 'in their own right', rather than in combination. BD and BDA are therefore no complementary resources (as proposed in Figure 1). These unexpected results will be discussed in section 5.

---

[7] SmartPLS Release 3 was used for the analysis (Ringle et al. 2015).

The structural model test statistics also reveal a not hypothesized but interesting effect: The control variable firm size has a significant *negative* direct ($\beta = -.190$, $p < .01$, $f = .048$*) and total ($\beta = -.16$, $p < .05$) effect on ABDM, which suggests that managers in smaller firms tend to base their decisions to a greater extent on analytics than their counterparts in larger firms.

Table 1 – Structural Model Results (Second Order Model)

n = 163	**Direct Effect (DE)**	**Indirect Effect (IE)**	**Total Effect (TE)**	***f* Square (DE)**
Model A				
BD Analytics → Performance (H1, H2)	*.138**	***.116****	***.254*****	*.019*
BD Analytics → AB Decision Making [H1, H2][*)]	.305***		.305***	.105*
AB Decision Making → Performance [H1, H2][*)]	.380***		.380***	.146*
Big Data → AB Decision Making (H3a)	.283***	***.123*****	.406***	.091
Big Data → Performance (H3a)		***.210*****	.210***	
Big Data → BD Analytics [H3][*)]	.404***		.404***	.198**
Size → AB Decision Making	*-.190***	*.030*	*-.160**	*.048**
Size → BD Analytics	*.097*		*.097*	*.011*
Size → Performance	*.109*	*-.048*	*.062*	*.014*
R Square				
AB Decision Making	.267***			
BD Analytics	.173***			
Performance	.207**			
Model B (Moderation):				
BD Analytics → Performance	.138*	.129**	.267***	.019
BD Analytics → AB Decision Making	.340***		.340***	.116*
AB Decision Making → Performance	.380***		.380***	.146*
Big Data → AB Decision Making	.254**		.392***	.068
Big Data → Performance		.205***	.205***	
Big Data → BD Analytics	.404***		.404***	.198**
Big Data*BD Analytics → AB Dec. Making (H3b)[#)]	***-.096***		*-.096*	*.010*
Big Data*BD Analytics → Performance		-.037	-.037	
Size → AB Decision Making	*-.202****	*.033*	*-.169***	*.054*
Size → BD Analytics	*.098*		*.098*	*.010*
Size → Performance	*.109*	*-.051*	*.059*	*.014*
R Square				
AB Decision Making	.274***			
BD Analytics	.173***			
Performance	.207**			

One-tailed: p<0.05; p<0.01**; p<0.001***; [*)] Part of indirect effect in H1, H2 and H3; [#)] Test for H3c redundant.*

5 Conclusion, Implications and Limitations

The research presented had three main objectives: The *first objective* was to empirically verify claims made primarily in the practitioner literature that Big Data (BD) and its analytics (BDA) leads to better performance and competitive advantage. The results for hypotheses 1 and 3a confirm that both BD and BDA have a positive effect on relative performance. In the case of BD, the effect is indirect via higher BDA sophistication and more analytics-based decision making (ABDM), whereas BDA affects performance directly *and* indirectly (via ABDM). These findings are important for both academia and industry. So far, evidence of benefits associated with BD and BDA was only case-based and not all cases reported were success-stories (LaValle et al. 2011, CGMA 2013). It is also obvious that BD/BDA initiatives come at costs, which are rarely disclosed when writing about 'success stories'. We related BD and BDA to four performance indicators which all measure organisational performance with reference to the firm's main competitor, so our findings suggest that BD and BDA can help outperforming competitors.

The *second* objective of our research was to verify to what extent any such benefits are achieved because BDA creates additional incentives for managers to base their decisions more on analytics (analytic-based decision making – ABDM). The test results for hypothesis 2 in conjunction with hypothesis 1 confirm a significant effect of BDA on performance via increased ABDM. That indirect effect is stronger than the direct effect, suggesting that the decision-impact of BDA is the main driver of performance (Barney 1986, Naor et al. 2008). There are of course many factors which co-determine to what extent managers base their decisions on analytics; firm size is one of those factors, and interestingly smaller firms base their decisions more on analytics than large firms. But to the best of our knowledge, no research has so far confirmed that higher BDA sophistication creates an incentive for managers to actually 'listen' to analytics and use it as base for their decisions. Further to that, we confirmed that using more analytics in decision making translates into significantly higher performance. The implications for practise are essentially that improving BDA actually influences decision making behaviour for the better, i.e. make it more evidence-based, which improves performance.

Finally, we intend to uncover any mediating and/or moderating effects of BD intensity on the above relationships. Our results suggest the presence of mediating rather than moderating effects. BD intensity is strongly related to BDA sophistication, confirming that using large and diverse data sets for analysis requires more sophisticated tools and methods, which in turn increases performance – directly and indirectly. So BD intensity in itself positively impacts performance, but only indirectly via the relationships described above. The absence of any moderating effects between BD and BDA on ABDM suggests that these resources influence decision making independently, i.e. the levels of BD and BDA do not have to be 'balanced' to achieve the desired outcomes.

As mentioned before, our study also reveals that managers in smaller organizations are significantly more likely to base their decisions on analytic outcomes than managers in large organizations. This finding is in line with some cases reported in the practitioner literature, which suggest that small businesses are in a good position to compete on analytics (CGMA 2013).

Like any study, our research is not free of limitations. Despite the fact that we deployed several procedural and statistical remedies to avoid biases (Podsakoff et al. 2003), survey-based research is never completely immune against biases. Second, the survey respondents were exclusively CIOs and other senior IT managers, a fact which inevitably introduces an IT-centric perspective. Future research could attempt to capture a more balanced perception, especially with regards to managerial decision making. Finally, we do not explicitly measure decision making quality or performance, but rather rely on prior research (Chaudhuri et al. 2011, Guillemette et al. 2014) which suggests that ABDM is associated with better decision making.

6 Appendix

Table 2 – Indicator and Constructs: Descriptive Statistics

	Mean	Std. Dev.	Skewness	Kurtosis	Skew/SE	Kurtosis /SE
Analytic Tools (BDA)						
Spreadsheets*)	6.55	0.795	-1.831	2.699	-9.637	7.140
BI Planning/Reporting Suites	4.88	2.056	-0.730	-0.703	-3.842	-1.860
Data ETL/Management Solutions	4.32	2.246	-0.263	-1.397	-1.384	-3.696
Statistical Suites – Basic Use	2.73	1.966	0.808	-0.665	4.253	-1.759
Statistical Suites – Advanced Use	2.42	1.866	1.104	-0.088	5.811	-0.233
Specialised Data Mining Suites	2.02	1.593	1.612	1.731	8.484	4.579
Data Visualisation Tools	3.53	2.215	0.143	-1.456	0.753	-3.852
BD/High Performance Computing Tools*)	2.13	1.709	1.358	0.708	7.147	1.873
Analytic Methods (BDA)						
Statistical Methods	3.44	2.114	0.300	-1.253	1.579	-3.315
Machine Learning, Data Mining, AI	2.45	1.846	0.965	-0.374	5.079	-0.989
OR, Optimisation Methods	2.53	1.789	0.828	-0.587	4.358	-1.553
Path Modelling*)	1.76	1.285	1.923	3.490	10.121	9.233
Volume (BD)						
Volume of Data – Access	4.51	0.781	-1.962	4.680	-10.326	12.381
Volume of Data – Use	4.13	0.972	-1.255	1.414	-6.605	3.741
Variety (BD)						
Diversity of Unstructured Data – Access	4.04	0.974	-0.885	0.214	-4.658	0.566
Diversity of Unstructured Data – Use	3.26	1.159	-0.266	-0.660	-1.400	-1.746
Velocity (BD)						
Rate of Change of Data – Access	4.16	0.831	-0.961	0.962	-5.058	2.545
Rate of Change of Data - Use	3.77	1.026	-0.676	-0.101	-3.558	-0.267
Analytic-Based Decision-Making						
Decisions about New Products/Services/Market	4.73	1.667	-0.697	-0.193	-3.668	-0.511
Decisions about Strategic/Key Suppliers	4.46	1.508	-0.557	-0.073	-2.932	-0.193
Decisions about Outsourcing/BPM	4.32	1.570	-0.542	-0.402	-2.853	-1.063
Decisions about Sales and Marketing	4.80	1.576	-0.561	-0.338	-2.953	-0.894
Decisions about Operations	5.02	1.486	-0.819	0.334	-4.311	0.884
Decisions about Procurement	4.52	1.446	-0.459	-0.137	-2.416	-0.362
Overall, Organisation Acts on Insights	4.78	1.445	-0.664	0.121	-3.495	0.320
Performance						
Sales Growth	5.13	1.194	-0.572	0.587	-3.011	1.553
Cost Reductions	4.72	1.147	-0.050	0.417	-0.263	1.103
Market Share	4.84	1.133	-0.352	0.617	-1.853	1.632
Profitability	4.94	1.304	-0.596	0.406	-3.137	1.074

Table 3 – VIF-Statistics

Independent Var. → Dependent Var.	BDA	BD	ABDM	PERF
Analytic Tools (BDA)	1.80			
Analytic Methods (BDA)	1.72			
Volume (BD)		2.07		
Variety (BD)		1.55		
Velocity (BD)		2.72		
Analytic-Based Decision Making			1.21	1.23
Big Data (BD)	1.20		1.20	
Performance				1.25

*) *Eliminated due to low loadings.*

Table 4 - Measures of Validity and Reliability (First Order Model)

	Composite Reliability	Cronbach's Alpha	Average Variance Extracted (AVE)	rho_A
Analytic Tools (BDA)	.861***	.807***	.511***	.825***
Analytic Methods (BDA)	.912***	.855***	.775***	.857***
Volume (BD)	.874***	.711***	.776***	.711***
Variety (BD)	.831***	.595***	.711***	.598***
Velocity (BD)	.866***	.691***	.763***	.696***
Analytic-Based Decision Making	.936***	.918***	.710***	.920***
Performance	.865***	.787***	.625***	.843***

One-tailed: p <.05; p <.01**; p <.001****

Table 5 - Harman's Single Factor Test

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.85	3.53	3.53	8.85	3.53	3.53
2	3.40	11.73	42.26	3.40	11.73	42.26
3	2.52	8.71	5.97	2.52	8.71	5.97
4	1.94	6.68	57.65	1.94	6.68	57.65
5	1.31	4.52	62.16	1.31	4.52	62.16
6	1.13	3.89	66.05	1.13	3.89	66.05

Table 6 – Fornell-Larcker Criterion for Discriminant Validity (First Order Model)

	Tools	Methods	Volume	Variety	Velocity	ABDM	PERF
Analytic Tools (BDA)	**.715**						
Analytic Methods (BDA)	.624	**.880**					
Volume (BD)	.320	.326	**.881**				
Variety (BD)	.300	.249	.378	**.843**			
Velocity (BD)	.320	.309	.716	.590	**.874**		
Analytic-Based Dec. Making	.346	.381	.385	.261	.359	**.843**	
Performance	.329	.184	.187	.003	.187	.418	**.790**

One-tailed: p <.05; p <.01**; p <.001****

Table 7 – Fornell-Larcker Criterion for Discriminant Validity (Second Order Model)

	BDA	BD	ABDM	PERF	SIZE
Big Data Analytics (BDA)	**1**				
Big Data (BD)	.405	**1**			
Analytic-Based Decision Making (ABDM)	.401	.405	**.843**		
Performance (PERF)	.301	.162	.418	**.79**	
Size	.098	.004	-.158	.063	**1**

Values in the diagonal are the square-roots of the AVE of each of the constructs.

Table 8 - Cross Loadings (First Order Model)

	Tools	Methods	Volume	Variety	Velocity	ABDM	PERF
Tool_2	**.541**	.195	.228	.099	.095	.288	.278
Tool_3	**.665**	.279	.266	.112	.160	.241	.289
Tool_4	**.778**	.603	.258	.296	.351	.302	.241
Tool_5	**.804**	.549	.234	.242	.250	.206	.231
Tool_6	**.767**	.521	.161	.248	.205	.297	.238
Tool_7	**.703**	.413	.254	.234	.257	.171	.174
Method_1	.495	**.880**	.289	.198	.263	.336	.161
Method_2	.555	**.865**	.289	.233	.257	.283	.159
Method_3	.594	**.895**	.282	.225	.294	.382	.165
BD_1	.209	.228	**.879**	.320	.651	.265	.133
BD_2	.354	.345	**.882**	.346	.611	.412	.196
BD_3	.201	.187	.263	**.828**	.516	.203	.003
BD_4	.301	.231	.370	**.858**	.481	.235	.003
BD_5	.170	.188	.590	.480	**.860**	.226	.145
BD_6	.379	.344	.659	.549	**.887**	.392	.180
ABDM_1	.332	.383	.383	.244	.352	**.863**	.348
ABDM_2	.352	.329	.318	.209	.291	**.872**	.351
ABDM_3	.263	.341	.276	.226	.316	**.863**	.318
ABDM_4	.229	.260	.370	.244	.341	**.836**	.377
ABDM_5	.290	.335	.283	.178	.262	**.827**	.382
ABDM_6	.280	.270	.306	.216	.245	**.792**	.335
PERF_1	.305	.172	.232	.106	.272	.365	**.850**
PERF_2	.194	.066	-.002	-.053	-.047	.190	**.493**
PERF_3	.241	.175	.188	-.011	.171	.368	**.865**
PERF_4	.294	.144	.115	-.058	.116	.362	**.886**

Table 9 - HTMT Values for Discriminant Validity (First Order Model)

	Tools	Methods	Volume	Variety	Velocity	ABDM
Analytic Methods (BDA)	.718					
Volume (BD)	.432	.418				
Variety (BD)	.412	.347	.577			
Velocity (BD)	.405	.395	1.02	.92		
Analytic-Based Dec. Making	.408	.427	.474	.351	.443	
Performance	.428	.217	.233	.113	.262	.484

7 References

Acito, F. and V. Khatri (2014). "Business Analytics: Why Now and What Next?" Business Horizons **57**(5): 565-570.

Bange, C. and N. Janoschek (2014) "Big Data Analytics 2014 - Towards a Data-Driven Economy."

Barney, J. B. (1986). "Organizational Culture: Can It Be a Source of Sustained Competitive Advantage?" Academy of Management Review **11**(3): 656-665.

Barské-Erdogan, A. (2014) "2014 Big Data Business Guide."

Baum, J. R. and S. Wally (2003). "Strategic Decision Speed and Firm Performance." Strategic Management Journal **24**(11): 1107-1129.

Brown-Liburd, H., H. Issa and D. Lombardi (2015). "Behavioral Implications of Big Data's Impact on Audit Judgment and Decision Making and Future Research Directions." Accounting Horizons **29**(2): 451-468.

Cao, L. (2016). "Data Science and Analytics: A New Era." International Journal of Data Science and Analytics **1**(1): 1-2.

CGMA (2013) "From Insight to Impact: Unlocking Opportunities in Big Data." CGMA Report. Chartered Global Management Accountant (CGMA), London.

Chakraborty, G., M. Pagolu and S. Garla (2013). Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS. SAS Institute Inc., Cary, North Carolina, USA.

Chaudhuri, S., U. Dayal and V. Narasayya (2011). "An Overview of Business Intelligence Technology." Communications of the ACM **54**(8): 88-98.

Chen, H., R. H. L. Chiang and V. C. Storey (2012). "Business Intelligence and Analytics: From Big Data to Big Impact." MIS Quarterly **36**(4): 1165-1188.

Davenport, T. H. (2006). "Competing on Analytics." Harvard Business Review **84**(1): 98-107.

Davenport, T. H. (2014). Big Data @ Work: Dispelling the Myths, Uncovering the Opportunities, Harvard Business Review Press, Boston, Massachusetts.

Davenport, T. H. and D. J. Patil (2012). "Data Scientist: The Sexiest Job of the 21st Century." Harvard Business Review **90**(10): 70-76.

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." MIS Quarterly **13**(3): 319-340.

Dehning, B. and V. J. Richardson (2002). "Returns on Investments in Information Technology: A Research Synthesis." Journal of Information Systems **16**(1): 7-30.

Deloitte (2014) "Technology in the Mid-Market Perspectives and Priorities."

Dhar, V. (2013). "Data Science and Prediction." Communications of the ACM **56**(12): 64-73.

Dillman, D. A., J. D. Smyth and L. M. Christian (2014). Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, John Wiley & Sons.

Dyché, J. (2014). " Technology for Big Data. " In Big Data @ Work: Dispelling the Myths, Uncovering the Opportunities. Editor T.H. Davenport, Harvard Business Review Press, Boston, Massachusetts.

Edwards, W. (1961). "Behavioral Decision Thoery." Annual Review of Psychology **12**(1): 473-498.

Fornell, C. and D. F. Larcker (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error." Journal of Marketing Research **18**(1): 39-50.

Garg, V. K., B. A. Walters and R. L. Priem (2003). "Chief Executive Scanning Emphases, Environmental Dynamism, and Manufacturing Firm Performance." Strategic Management Journal **24**(8): 725-744.

Gartner Research (2015) "Survey Analysis: Practical Challenges Mount as Big Data Moves to Mainstream."

Guillemette, M. G., M. Laroche and J. Cadieux (2014). "Defining Decision Making Process Performance: Conceptualization and Validation of an Index." Information & Management **51**(6): 618-626.

Hagel, J. (2015). "Bringing Analytics to Life." Journal of Accountancy **219**(2): 24.

Hair Jr, J. F., G. T. M. Hult, C. Ringle and M. Sarstedt (2014). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), Sage Publications.

Hayes, A. F. (2015). "An Index and Test of Linear Moderated Mediation." Multivariate Behavioral Research **50**: 1-22.

Henseler, J., C. M. Ringle and M. Sarstedt (2015). "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling." Journal of the Academy of Marketing Science **43**(1): 115-135.

Holsapple, C., A. Lee-Post and R. Pakath (2014). "A Unified Foundation for Business Analytics." Decision Support Systems **64**: 130-141.

IEEE (2013) "Scenarios for IEEE." Envisioning Our Future. Available online at https://www.computer.org/documents/8856435/47968494/IEEE+world+scenarios+of+2030/e816ba5a-78e1-4743-a24e-2acdf3bb6611;jsessionid=73230b0f4032b2f1cb3fc7894294; accessed 14/5/2017.

King, G. (2013). Big Data Is Not About the Data! Golden Seeds Innovation Summit, New York City.

Kraska-Miller, M. (2014). Nonparametric Statistics for Social and Behavioral Sciences, CRC Press.

LaValle, S., E. Lesser, R. Shockley, M. S. Hopkins and N. Kruschwitz (2011). "Big Data, Analytics and the Path from Insights to Value." MIT Sloan Management Review **52**(2): 21-32.

Linoff, G. and M. J. A. Berry (2001). Mining the Web : Transforming Customer Data into Customer Value. New York, John Wiley & Sons.

March, J. G. (1978). "Bounded Rationality, Ambiguity, and the Engineering of Choice." The Bell Journal of Economics **9**(2): 587-608.

Masli, A., V. J. Richardson, J. M. Sanchez and R. E. Smith (2011). "The Business Value of It: A Synthesis and Framework of Archival Research." Journal of Information Systems **25**(2): 81-116.

McAfee, A. and E. Brynjolfsson (2012). "Big Data: The Management Revolution." Harvard Business Review **90**(10): 60-68.

Meissner, P. and T. Wulf (2014). "Antecendents and Effects of Decision Comprehensiveness: The Role of Decision Quality and Perceived Uncertainty." European Management Journal **32**(4): 625-635.

Melville, N., K. Kraemer and V. Gurbaxani (2004). "Information Technology and Organizational Performance: An Integrative Model of IT Business Value." MIS Quarterly **28**(2): 283-322.

Minelli, M., M. Chambers and A. Dhiraj (2012). Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, John Wiley & Sons.

Mintzberg, H. and J. A. Waters (1982). "Tracking Strategy in an Entrepreneurial Firm." Academy of Management Journal **25**(3): 465-499.

Moffitt, K. C. and M. A. Vasarhelyi (2013). "AIS in an Age of Big Data." Journal of Information Systems **27**(2): 1-19.

Naor, M., S. M. Goldstein, K. W. Linderman and R. G. Schroeder (2008). "The Role of Culture as Driver of Quality Management and Performance: Infrastructure Versus Core Quality Practices." Decision Sciences **39**(4): 671-702.

Nowduri, S. (2011). "Management Information Systems and Business Decision Making: Review, Analysis, and Recommendations." Journal of Management and Marketing Research **7**: 1.

Parmar, R., I. Mackenzie, D. Cohn and D. Gann (2014). "The New Patterns of Innovation." Harvard Business Review **92**(1/2): 86-95.

Peters, M. D., B. Wieder, S. G. Sutton and J. Wakefield (2016). "Business Intelligence Systems Use in Performance Measurement Capabilities: Implications for Enhanced Competitive Advantage." International Journal of Accounting Information Systems **21**: 1-17.

Pfeffer, J. and R. I. Sutton (2006). "Evidence-Based Management." Harvard Business Review **84**(1): 62-74.

Podsakoff, P. M., S. B. MacKenzie, J.-Y. Lee and N. P. Podsakoff (2003). "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies." Journal of Applied Psychology **88**(5): 879.

Preacher, K. J., D. D. Rucker and A. F. Hayes (2007). "Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions." Multivariate Behavioral Research **42**(1): 185-227.

Rindfleisch, A., A. J. Malter, S. Ganesan and C. Moorman (2008). "Cross-Sectional Versus Longitudinal Survey Research: Concepts, Findings, and Guidelines." Journal of Marketing Research **45**(3): 261-279.

Ringle, C. M., S. Wende and J.-M. Becker (2015). "Smartpls 3. Bönningstedt: SmartPLS. Retrieved from http://www.smartpls.com.

Rossino, A. (2015) "Federal Spending on Cloud and Big Data Are on the Rise."

Rousseau, D. M. (2006). "Is There Such a Thing as "Evidence-Based Management"?" Academy of Management Review **31**(2): 256-269.

Schroeck, M., R. Shockley, J. Smart, D. Romero-Morales and P. Tufano (2012) "Analytics: The Real-World Use of Big Data."

Simon, H. A. (1972). "Theories of Bounded Rationality." Decision and Organization **1**(1): 161-176.

Turban, E., D. Leidner, E. McLean and J. Wetherbe (2008). Information Technology for Management : Transforming Organizations in the Digital Economy, Wiley.

Vasarhelyi, M. A., A. Kogan and B. M. Tuttle (2015). "Big Data in Accounting: An Overview." Accounting Horizons **29**(2): 381-396.

Watzlawick, P. (1984). Invented Reality: How Do We Know What We Believe We Know?

Wieder, B. and M.-L. Ossimitz (2015). "The Impact of Business Intelligence on the Quality of Decision Making – a Mediation Model." Procedia Computer Science **64**: 1163-1171.

Yang, C. C., H. Chen and K. Hong (2003). "Visualization of Large Category Map for Internet Browsing." Decision Support Systems **35**(1): 89-102.