5-18-2013

A REALIST EXAMINATION OF BUSINESS ANALYTICS

Michael J. Cuellar
North Carolina Central University, mcuellar@georgiasouthern.edu

A REALIST EXAMINATION OF BUSINESS ANALYTICS

Michael J. Cuellar
North Carolina Central University
mcuellar@nccu.edu

Second author's na
Affiliation
e-mail address

ABSTRACT

Business analytics has become the latest fad in management practice. Carried out by largely positivist statistical analysis, anecdotal evidence claims great success for these practices. This paper takes a critical realist approach to a philosophical analysis of these practices. Reviewing CR ontology and epistemology, it applies those to the practice of business analytics, showing that where BA success occurs it is because the analysis has encountered relatively unimpeded actions by relatively enduring structures of causal mechanisms. It may fail in those areas undergoing rapid structural change, or where there exists a confusing welter of mechanisms. The danger in this approach from a CR perspective is that other causal mechanisms may intrude, or the structure of the mechanisms may change causing the models built to fail. This paper argues that the practice of BA may be improved by performing retroductive analysis to identify these structures so that anticipatory action can be taken to avoid failure of the models.

Keywords

Business Analytics, Critical Realism, Induction, Retroduction, Morphogenesis

INTRODUCTION

Business Analytics, "the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to derive decisions and actions" (Davenport et al. 2007), has attracted more and more interest in the practitioner community. Driven by success stories such as those of Harrah's Entertainment and First American Bank where customer analytics that allowed them to improve top-side performance and by reports that top performing companies more likely to use analytics than not and companies using analytics get a 5-6% improvement in performance (Watson 2012), the adoption of business analytics has continued to grow. The Q4 2009 Forrester survey shows that, where reporting tools are concerned, 74 percent of enterprises either already use them or planned to implement them in 2010. For predictive analytics and data mining (PA/DM), the equivalent number is 31 percent; for text analytics and complex event processing it's 19 percent; and for in-database analytics, 7 percent(Kobielus 2010). Similarly, TechTarget reported that despite tighter IT budgets because of the weak economy, companies still invested in BI tools in 2010. More respondents said their organizations adopted BI software: the vast majority of businesses – 85% – have at least one BI tool in use. More than half of the organizations represented were using multiple BI tools: 52%, same as in 2009 (Aucoin 2010).

The predominant form of BA processing is the inductive statistical analysis of operational data. In this analysis, data, often from the company's operational systems is analyzed by various forms of statistical analysis to identify patterns of behavior so that valuable predictions can be made to optimize the company's performance. One of the emphases of the use of BA is to move from management by hunch and intuition to management by facts and data. And not only that, the direction of the technology according to SAS is to move beyond reactive, reporting and analysis to provide "Web-service-enabled analytics or capabilities that enable operational applications." (Troester 2012,p. 9).

Given this increased interest in the area, it is perhaps appropriate to consider the basis on which this activity is based. What is it in the current practice that gives this area its warrant to claims that is provides useful information to practice? Philosophically, BA relies upon positivist meta-theory. This meta-theory is severely empirical. The data extracted from operational systems is utilized often without interpretation or attempts at explanation of causal factors to create the models on which decisions are based. Patterns in the data are statistically determined and given the force of law like relations and used to analyze other data to make decisions. This positivism runs directly into the problem of induction: the idea that we can generalize from particulars to law-like generalities or from past actions to future occurrences. In general, from Hume on it has been held that this type of logic is problematic. It has been expressed in the popular culture in the maxim: "past results are no indication of future performance." This warning is seen on many prospectuses of financial investments and caveats on election predictions.

This paper examines the philosophical basis for Business Analytics from a Critical Realist perspective. First, we look at the problem of induction and discuss what it is and its roots within the implicitly flat Humean empiricist ontology and show how it fails to explain that fact that BA has both successes and failures. It also fails to provide a discussion of where it will succeed and where it will fail. Following that, we will discuss the alternative of Critical Realism. We will show that the structured explicit ontology of Critical Realism provides explanatory purchase on explaining why BA succeeds and why it fails and give us some direction as to how to improve the practice of BA in order to achieve higher reliability. We illustrate the philosophical argument with an illustrative example of using regression to induce a model and then showing how that empiricism fails to provide the right model and that CR can provide direction to improve the model.

OVERVIEW OF BA

By BA, we refer to inductive statistical analysis of quantitative data."Data mining is a largely automated process that uses statistical analyses to sift through massive data sets to detect useful, non-obvious, and previously unknown patterns or data trends. The emphasis is on the computer-based exploration of previously uncharted relationships By uncovering such previously unknown relationships, managers have the potential to develop a winning marketing strategy that increases their hotel's bottom line." (Magnini et al. 2003). Typical applications for BA include such things as customer acquisition, retention, upselling, churn, and pricing tolerance as well as optimization of staffing, inventory, deliveries, financial forecasting, insurance rate setting, and fraud detection. It is contraindicated where a decision is needed immediately and there is no time to gather data, where no data is available on the problem, where the situation has changed such that the past is not applicable, when the event of interest so rarely occurs enough data is available, where you can't measure the variables and where we are looking for expert guidance through a decision process (Davenport et al. 2010; Truxillo 2012). The statistical techniques used include: descriptive statistics, genetic algorithms, neural networks, cluster analysis, association rules, linear and logistic regression, and decision trees(Magnini et al. 2003; Truxillo 2012). Success factors include having accessible high-quality data, an enterprise-wide orientation, leadership, strategic targets to aim toward and good analysts (Davenport et al. 2010).

BA METHODOLOGY

There is a process to BA. First and most important a business objective or question is established. Then the data needed to answer the question is identified and extracted from operational databases, or collected via surveys. This data is then transformed and cleansed in order to be in the form required for analysis. Often this includes interpolation of missing data by various statistical means. The data is then divided into "training", "validation" and sometimes "test" data sets. Then a series of models is built from the training data. These models are such things as structured equations or decision trees that can be used to analyze data related to the entities under analysis. Model selection follows the concept of "honest assessment" (Truxillo 2012). The object is to select the model that performs best on the validation dataset. This selection is a tradeoff between bias and variance. If the model which is the best fit to the training dataset (i.e. highest R squared value) is selected this could result in "overfitting" the data – accommodating the nuances of the training dataset (noise, high variance). Simply guessing could result in an under fitted model which misses the signal(high bias). Therefore, the model that performs best on the validation dataset is the one that is to be selected. This model can then be used on freshly collected data to predict behavior of the entities under analysis. If a "test" dataset has been created, it is used to measure prediction accuracy of the model (Truxillo 2012).

BA thus follows a positivist meta-theory. The operational, survey and other data that we analyze areheld to be sufficient to give us clear access to reality. The statistical regularities that we develop are models of consumer and other entities behavior that can be used for prediction. In empiricist thought, a law is a statement of universal form. Statements about events: If A occurs then B occurs. Sometimes these are probabilistic. The question arises then: how do we know if it is true and not a spurious correlation?Hempel's answer is if it supports counterfactual or subjunctive conditionals or can serve as an explanation(Hempel 1966). Philosophically stated, BA uses induction to identify universal laws that can be used to guide decision making. This leads us directly to the problem of induction.

THE PROBLEM OF INDUCTION

Simply stated, the problem of induction is the question of whether we have a warrant to move from specific instances to general statements (induction proper) or from past observations to future events (eduction). The issue here is how we can be assured that this pattern, which we observe in the past, will continue into the future. First proposed by Hume (1999), no satisfactory answer has been given to this problem. Hume argued that our knowledge of cause and effect is not learned a priori but through experience. Further, every effect is distinct from its cause and is not to be found in the cause. We cannot therefore discover why an event should be the cause of another as knowledge is not founded on reasoning but on experience which causality is not present. Finally, our supposed knowledge of cause and effect is founded upon our supposition that the future will be conformable to the past. Nature however can change and therefore the past can be useless as a predictor of the future (Hume 1999). Russell (1912) similarly states that in order for induction to hold there must be a "principle of induction" which states a) that when event A has been found in the presence of event B and when more cases of this are found, the greater the probability is greater that they will be associated in a new case; and b) that a sufficient number of cases of association will make the probability of association a near certainty. This principle can neither be proved nor disproved by experience. Thus, all statements made on the basis of experience, which statements attempt to tell us something about what we have not experienced, is based itself on a principle which itself cannot be proven.

There have been many attempted challenges and rehabilitations to the problem of induction offered. For example, Popper, while accepting Hume's analysis, attempted to show that science really didn't progress by means of induction but rather deduction (Bhaskar 1997). Other attempts have been make to justify induction inductively, change inductive arguments to probabilitic statements or to dissolve the problem (Bhaskar 1997). Edwards (1949) attempts to argue contra Russell that this is simply philosophical pettifoggery: we know perfectly well that a man jumping off the Empire State building will fall toward the ground because in every instance where humans have been unsupported from an elevated position, they have fallen down. Strawson(1989) in his application of Wittgenstein's concept of language games and while not denying the philosophical truth of the problem of induction, argues that the question, while not senseless, is idle. Idle in the sense of the fact that we have our inescapable commitments that we will not give up.

From a Critical Realist standpoint, the problem of induction stems from the faulty Humean ontology employed in all empiricist meta-theory. As Hume suggests, according to this meta-theory, the only way to know what exists is via experience. That is we cannot reason to the nature of cause and effect *a priori*, we must rather deal with it through our senses. This principle reduces what exists to what may be known. It results in an implicit, flat ontology of discrete, atomistic events. These events have no necessary connection with each other and therefore can offer no justification for assuming that we can at anytime gain knowledge by inducing general states from particulars or knowledge of the future from the past.

In the next section, we will summarize the meta-theoretic commitments of critical realism and then show how CR can resolve the problem of induction and then provide a basis for understanding how business analytics can work.

CRITICAL REALIST METATHEORY

There are many excellent introductions to Critical Realist philosophy (Ackroyd et al. 2000; Mingers 2004; Mutch 2007) and they are not repeated here. This paper does provide a discussion of key concepts necessary to understanding how CR meta-theory addresses the problem of induction and how it conceives of the project of BA. This section therefore covers BA meta-theory: ontology and epistemology, and the nature of social entities vis-à-vis natural entities.

Ontology

CR is a realist philosophy of science. That means that CR holds that there exists a real world of objects that exist separately of man and his conception of it. This is not unique. While interpretivism rejects a realist view in favor of one where the world is a social construction, positivism also holds to a realist position. The hallmark of CR that differentiates it from positivism is its structured and layered ontology. In contrast to the implicit, flat ontology of discrete events posited by empiricism, CR explicitly holds to a structured three-part ontology (Figure 1).

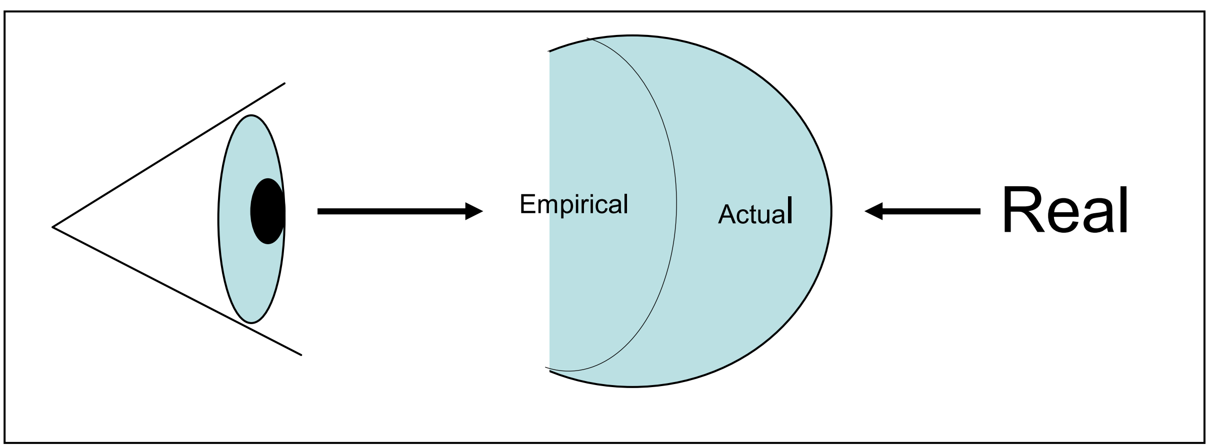

FIGURE 1: Critical Realist View of Ontology

The realm that we experience is the empirical. We experience the pencil falling or the pencil being held in our fingers. We feel the force of gravity as we stand or sit. The empirical is a subset of events that actually occur, the realm of the actual. The subset is of those events that we can actually experience. The actual is the realm of all the events that occur whether or not a human senses them. There are sound waves below and above out ability to hear them. Gravity still exists even if a pencil doesn't fall. If a tree falls in a forest with no one to hear it, it still makes a noise. These events are caused by properties inherent in structures of real things. This is the realm of the real. The earth is a structure that has properties such as gravitational force. For Bhaskar, a thing is real if it is capable of causing events in the Actual realm (Bhaskar, 1997). The events that occur in the actual and empirical realms are caused by the structure and properties of the objects in the Real realm. These properties may or may not manifest effects in the other realms since they may not always be in operation or visible or may be over ruled by properties of other objects.

Epistemology

As noted above, CR holds that we only have direct access to the empirical realm. Our finitude creates an inability to know about every event that exists in the actual realm or every entity in the real one. To gain a conception of the Real structures and mechanisms that cause the observed phenomena, we must perform *transcendental analysis*. This consists of asking the question as to what structures and mechanisms must exist to cause these phenomena. We *transcend* the empirical facts to arrive at that which is not empirically sensible but required for the phenomena to exist. While this bears a resemblance to interpretive methods such as phenomenology, it is different in that it is based on the structured ontology described above as opposed to the flat ontology proposed for interpretivism and a different conception of generalization. The concept of generalization employed by positivism or interpretivism is one of extent; seeking to expand the range of areas where the rule applies. In CR, generalization is vertical; it seeks to penetrate the layers of phenomena to arrive at the basic structures and mechanisms that cause them. The critical realist concept of science is thus an ongoing social activity with the aim of production of knowledge about the mechanisms about production of the phenomena in nature that combine to produce the flux of the phenomena of the world. It is the systematic attempt to express in thought the structure and ways of acting of things that exist and act independently of human thought (Bhaskar 1997). The basic process of scientific discovery is a dialectic. A regularity is discovered, a number of attempts are made to build models that explain the regularity by means of generative mechanism. These in turn are subjected to empirical test.

Thus our knowledge of the real is mediated by means of theories. We know things only by the interpretation that our concepts give phenomena. It is a social product constructed from previously derived social products that point to objects in the realm of the real. We see that knowledge is not only produced but is in an ongoing process of transformation. We see then that knowledge may change independently of the Real objects or the Real objects may change independently of knowledge. There is no necessary connection between them.

Critical Realism and Induction

The critical realist ontology and epistemology provide an answer to the problem of induction. The problem occurs because in Humean meta-theory, there is no necessary connection between events, thus there is no means by which to say that a regularity observed in the past will continue to the present. CR resolves the problem by identifying the mechanisms responsible for bringing a phenomenon to pass. We can say that phenomenon B occurs because entity x has a propensity to do A under conditions C_1, $C_{2,\ \ldots}C_n$ because of its nature N (Bhaskar 1997). Through its stratified ontology, CR has the power to create the connection between events missing in Humean meta-theory.

Now, as said above, we do not have direct knowledge of real. The warrant to assume that E_2 will continue to follow E_1 is because we have 1) a model of the nature of the phenomena that describes the real realm in such a way that when E_1 occurs under conditions C_1, $C_{2,\ldots\ldots}C_n$, E_2 must occur because of natural necessity; and 2) that this model is has the highest explanatory capability of all competing models if the mechanisms postulated to cause the events are not empirically observable.

BUSINESS ANALYTICS AND CRITICAL REALISM

In this section, we provide a critical realist examination of the practice of business analytics as conducted by inductive statistical analysis. The structured ontology is key to the understanding of how BA may be successful. When BA does its inductive statistical analysis, it only has access to the Empirical realm. This limited access deprives BA of certainty in the success of its efforts to model the behavior of its subjects and find predictive accuracy with its results.

From a critical realist perspective, we can make the following observations about this practice:

1. The data used by BA is a record of events that occurred at the empirical level.

For example, the data might consist of records that show a customer with certain characteristics (male/female, lives in a certain zip code, age) bought or didn't buy a particular item (Figure 2). This data would be collected from a point of sale or other system and augmented from other data such census data, which provides other information. This data arises as a result of customer interactions with the company, e.g. inquiries, sales, purchases, etc. Thus what is available to BA is only historical records of empirical events. What BA knows about the entities is the empirical characteristics of them – it does not know the structure or causal mechanisms of the entity.

Empirical Level	ID	Item	Gender	Age	Postal Code	Etc. …..

FIGURE 2: The Data Elements in the Empirical Level

2. The event occurred because of a network of relationships of entities and causal mechanisms in a particular environment that responded in that way under a set of stimuli

In CR meta-theory, events occur as causal mechanisms are activated. Thus, for the transaction data to be recorded, a certain set of causal mechanisms from certain entities under certain conditions received stimuli, which resulted in the transaction that left the data analyzed by the BA system. From the CR ontology, we understand that the all events occur as the result of the action of some causal mechanism or mechanisms. However as the statistical analysis is based on data collected in the empirical realm, it does not penetrate to the level of the real and therefore knows nothing really about the causal mechanisms that bring the regularity to pass (Figure 3).

As figure 3 shows, the entities in the real realm interact under stimulation. The nature of these entities and their relationships cause the particular data that BA finds at in the empirical realm.

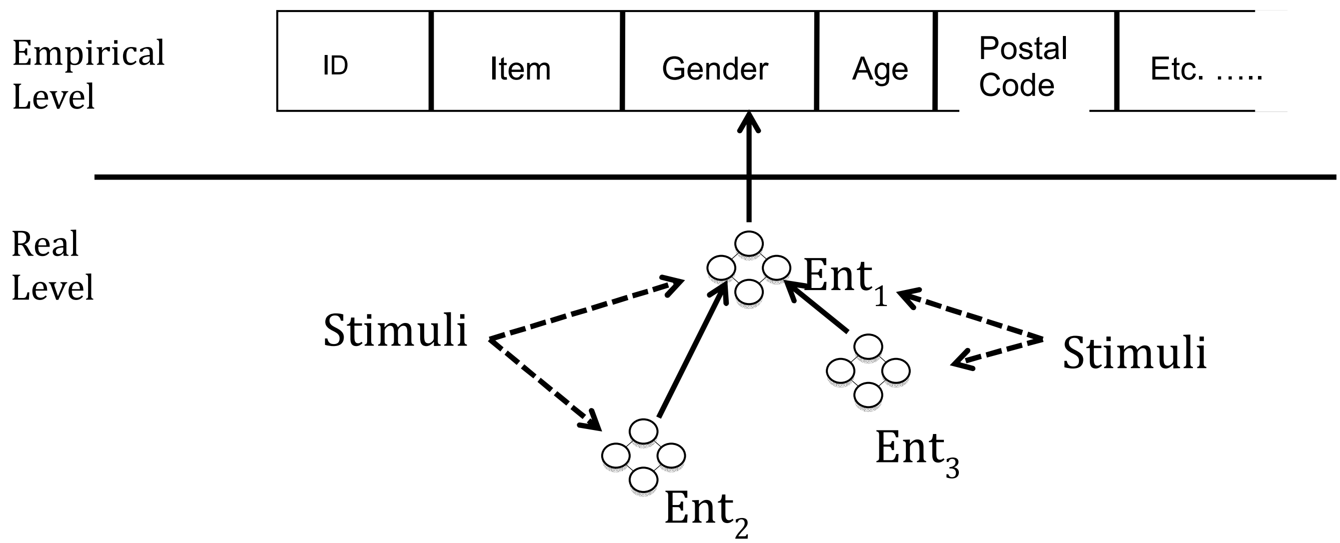

FIGURE 3: The Entities in the Real Realm Producing the Empirical Effects

3. BA identifies relationships such as correlation or covariance between certain factors and the end result under study. What BA presents then is regularities at the empirical level in open systems.

What we learn is e.g. how the characteristics of the customer contributed to the end result. Using logistic regression, we might learn that a customer of a certain gender, living in a certain postal code would have an X% probability of purchasing this particular item when using certain advertisements. We learn nothing about causal conditions or boundary limitations but simply what the probability of the desired behavior is.

4. To identify this correlation requires that sufficient instances of the correlation exist to be detectable as different from the norm by the statistical analysis.

In order to be detected by statistical analysis, a regularity must exist in enough observations to be significantly different than the average response. While BA does not require statistical significance, in order to show up in the statistical analysis, sufficient numbers of observation showing the regularity must exist in the sample. Thus as shown in figure 4, repeated instances of the data produced by the real entities must appear in the sample. There must have been repeated interactions of the entities under the stimuli to produce these instances.

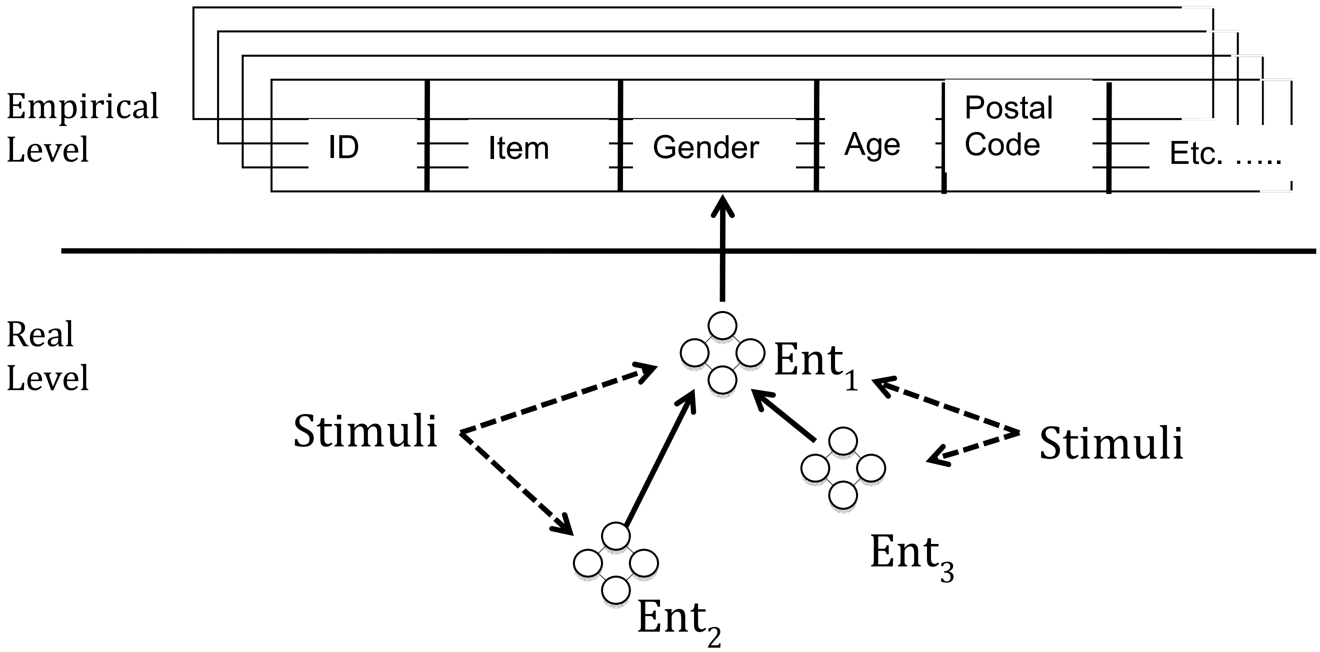

FIGURE 4: The Stability of the Real Entities Producing Consistent Empirical Results

5. In order for the regularity to exist, the entities, mechanisms and stimuli must have been in a relatively stable configuration for enough occurrences of the regularity to be identified as different from the background "noise" of the rest of the environment.

This is a critical point. The relationship between the entities, the key social structures and the stimuli must be relatively stable to produce the data required. The effects of the actions also must not be obscured by the actions of other entities on the entities of the regularity or entities that would cause effects that would hide the effects of the entities in the regularity. This regularity of behavior will continue as long as the relatively stable configuration of entities and stimuli continues to exist.

6. Thus BA's ability to predict depends upon the stability of the relationship between the entities, the entities internal structure and the nature of the stimuli occurring external to the entities.

This point is what makes BA successful or a failure. If this relationship changes or if the effects of the structure are interfered with or obscured by other mechanisms, then the model built based on past information will fail.

DISCUSSION AND CONCLUSIONS

Business Analytics is an emerging topic of interest in the practitioner community. Anecdotal stories of tremendous success have been raised for BA. The question that was considered was if these claims are justified and if so what is the limit of that justification. We saw that existing BA is based on positivist meta-theory using inductive statistical analysis. This analysis led us up against the problem of induction: the question of how can we be warranted to move from particular instances to general statements or from past actions to predictions of the future. We saw that empiricist philosophical analysis could not provide any such warrant nor could it explain either success or failure of the BA enterprise. This was due to Humean implicit ontology of discrete atomistic events which have no necessary connection with each other.

It was suggested that Critical Realism could fulfill that gap. CR rejects the Humean ontology in favor of a structured ontology with a realm of real things with causal mechanisms that bring Actual events to pass of which we can observe some in the Empirical realm. This ontology indicates that BA works with artifacts of the actions of entities left in the Empirical realm. Its ability to spot these actions depends on the relatively enduring relationships between the entities, the environment that they are in and the stimuli received. We saw in a simple example that BA in the form of SAS Enterprise Miner ™ is able to identify the relationships.

In this concluding section, we would like to propose some concepts about the applicability of BA derived from the CR ontology:

Proposition 1: Business Analytics will be more successful if the field which the BA project investigates is smaller (has less entities in it).

As shown in our example, if the field under examination has only three entities in it, it is easy to identify the effects between them. Adding more entities and interactions makes it more difficult for BA to identify the appropriate relationships and their effects. This would account for the failure of some statistical analyses, e.g. econometrics (Lawson 1997; Mingers 2006). While it is not possible to close the field and eliminate entities from it, BA practitioners should use their domain knowledge to assess which are the most important entities in the field to be considered.

Proposition 2: The models generated by Business Analytics must be continuously reviewed to ensure that they accurately reflect the regularities in the field studied.

As we have seen, if the relationships between entities, their environment and stimuli change, the ability of models derived will change. It is necessary then, that practitioners continuously or on a regular basis evaluate the performance of the models or redevelop their models based on current data to ensure that the relationships between the entities are appropriately modeled.

Proposition 3: Business Analytics will be more successful if there is an opportunity to investigate the causal mechanism by means of retroductive analysis.

It is important to know how stable this configuration will be and how long will it continue to exist. Additionally, it would be good to know what events will cause it to break apart. CR provides a mechanism to understand these points. As described above, given the structured ontology, it is necessary to transcend empirical facts to arrive at a theory of the relationship of entities and causal mechanisms that bring these facts to pass. While it is beyond the scope of this paper to describe in full these methods, many works provide these details (Cuellar 2010; Danermark et al. 2002; Sayer 1992). What follows is a brief description of this methodology.

Step 1: Effectively describe the problem. Here descriptive analysis is done to describe and define what the phenomenon under study is and under what conditions it occurs. Descriptive statistics and business analytics is useful for this activity (Finch et al. 2002; Lawson 1997).

Step 2: Abduct possible structures of causal mechanisms based on exist theories.In this step, existing theories of the phenomenon are consulted and the phenomenon is redescribed in terms of those theories. Where the theory is not derived from a CR meta-theoretical base, it must be redescribed into CR terms. This creates a set of possible explanations of the phenomenon in terms of existing theories. Theories may also be combined and modified in this process.

Step 3: Retroduction of new theory. Where existing theories are non-existent, ineffective or weak, new theory is developed. This is done by asking the question “what must reality be like in order for this phenomenon to occur?” Here novel descriptions of entities and causal mechanisms are derived to explain the phenomenon.

Step 4: Selection of the Explanatory Theory. In this step, the theories are analyzed to determine which one explains the phenomenon most effectively; which has the highest explanatory power. Where possible the theory is examined against related phenomena. In the case of BA, we would ask the question, which theory provides the best explanation for the statistical results received?

Step 5: Concretization, Here the selected theory is used to provide an explanation for the phenomenon observed, in this case the BA results. The object here is to explain how the results derived from the theory. For use by BA practitioners, this would allow them to understand the relationship of the entities and their causal mechanisms and give insight to them on how to change their practices to be more effective in improving their businesses around these entities.

Once derived, this theory can be further subjected to CR based Morphogenetic analysis which analyzes the nature of the social structures involved to determine where the fault lines for change might exist (Archer 1995; Cuellar 2010). Identification of the fault lines will allow the practitioner to understand what events might occur that might cause realignment or the entities and thus cause their methods to become invalid.

The conclusion of this study is that from a Critical Realist point of view, Business Analytics is a method to identify the effects of interacting social entities. This interaction is subject to change based on changes to the relationships between the social entities. Retroductive analysis can assist practitioners in identifying the relationships of the structures and morphogenetic analysis can assist in identification of stress points in social structures for change and the events that might cause them to change. This will allow BA practitioners to become more accurate with their models and avoid model failure in a timely manner. In the short space of this paper, we did not consider the effects of the open nature of social science in any detail. Nor was the practical aspect of performing these analyses within a business setting. This philosophically based study provides a philosophical base for the practice of Business Analytics. Further research will be needed to further develop these insights and to improve practice based on these propositions.

REFERENCES

1. Ackroyd, S., and Fleetwood, S. "Realism in Contemporary Organization and Management Studies," in: *Realist Perspectives on Management and Organizations,* S. Ackroyd and S. Fleetwood (eds.), Routledge, London, 2000, pp. 3-25.
2. Archer, M. S. *Realist Social Theory: the morphogenetic approach*, (1st ed.) Cambridge University Press, Cambridge, 1995, p. 354.
3. Aucoin, J. "2010 business intelligence (bi) challenges and priorities survey," TechTarget, 2010.
4. Bhaskar, R. *A Realist Theory of Science*, (2nd ed.) Verso, London, 1997, p. 284.
5. Cuellar, M. J. "Using Realist Social Theory to Explain Project Outcomes," *International Journal on IT Project Management* (1:4) 2010, pp 38-51.
6. Danermark, B., Ekstrom, M., Jakobsen, L., and Karlsson, J. C. *Explaining Society Critical Realism in the Social Sciences*, (1st ed.) Routledge, London, 2002, p. 221.
7. Davenport, T. H., and Harris, J. G. "The Dark Side of Customer Analytics," *Harvard Business Review*), May 2007.
8. Davenport, T. H., Harris, J. G., and Morison, R. *Analytics at Work Smarter Decisions Better Results* Harvard Business School Publishing Corporation, Cambridge, MA, 2010.
9. Edwards, P. "Russell's Doubts about Induction," *Mind* (68) 1949, pp 141-164.
10. Finch, J. H., and McMaster, R. "On categorical variables and non-parametric statistical inference in the pursuit of causal explanations," *Cambridge Journal of Economics* (26) 2002, pp 753-772.
11. Hempel, C. G. *Philosophy of Natural Science* Prentice-Hall, Inc., Englewood Cliffs, N. J., 1966, p. 116.
12. Hume, D. "Sceptical Doubts Concerning the Operations of the Understanding," in: *Readings in Epistemology,* I. Jack S. Crumley (ed.), Mayfield Publishing Company, Mountain View, California, 1999.
13. Kobielus, J. "Number of People Using Advanced Analytics," in: *James Kobielus' Blog for Business Process Professionals*, 2010.
14. Lawson, T. *Economics and Reality* Routledge, London and New York, 1997, p. 364.
15. Magnini, V. P., Honeycutt, J., Earl D., and Hodge, S. K. "Data Mining for Hotel Firms: Use and Limitations," *Cornell Hotel and Restaurant Administration Quarterly* (44) 2003.
16. Mingers, J. "Real-izing information systems: critical realism as an underpinning philosophy for information systems," *Information and Organization* (14) 2004, pp 87-103.
17. Mingers, J. "A critique of statistical modelling in management science from a critical realist perspective: its role within multimethodology," *Journal of the Operational Research Society* (57:2), Feb 2006, pp 202-219.
18. Mutch, A. "Concerns with "Mutual Constitution": A Critical Realist Commentary," in: *Issues and Trends in Technology and Human Interaction,* B.C. Stahl (ed.), IRM Press, Hershey, PA, 2007, pp. 230-244.
19. Russell, B. "On Induction," in: *The Problems of Philosophy*, Oxford University Press, 1912.
20. Sayer, A. *Method in Social Science A Realist Approach*, (2nd ed.) Rutledge, London, 1992, p. 313.
21. Strawson, P. F. "Hume, Wittgenstein and Induction," in: *Knowledge and Scepticism,* M. Clay and K. Lehrer (eds.), Westview Press, 1989, pp. 31-49.
22. Troester, M. "Your top 10 checklist for IT," *sascom*:second quarter) 2012.
23. Truxillo, C. *Advanced Business Analytics* SAS Institute, Cary, NC, 2012.
24. Watson, H. "BI and How I teach it," in: *Southern Association for Information Systems Conference* SAIS, Atlanta, Georgia, 2012.