ABSTRACT

Big data is often seen as a silver bullet when applied with the appropriate data analytics tools; however, there is a downside: dependence. Over time, data analytics users begin to apply less cognitive effort or apply data analytics in the wrong scenario, which can have a long-term negative effect on firms. The limits to the applicability of big data has not been vetted in data analytics research. Through General Systems Theory and causal map modeling, this work provides a perspective on a negative of big data analytics and conceptually adds to the practical and theoretical understanding of big data analytics.

KEYWORDS: Data Analytics, General Systems Theory, Technology Dependence, Cognitive Effort

INTRODUCTION

The potential of big data and business analytics is seemingly limitless. Since entering the list, big data is consistently in the top ten of most significant investments by Information Technology leaders (Kappelman, McLean, Johnson, & Gerhart, 2014). While most companies see big data as an opportunity, it does not instantly and easily solve business problems unless properly handled. By one estimate, the amount of data doubles every year (Helbing, 2014). While big data is leveraged in many businesses, the research on how analysts are actually using big data and improving strategy is lagging.

Big data is not just a large quantity of data, instead it incorporates the 3Vs: volume, velocity, and variety (Gartner Inc., 2013). The combination addresses the amount, the speed it is collected, and the many sources and structures of the data. Big data alone can be a problem more than a solution from a strategic perspective; creating value out of big data requires big data analytics. Big data analytics has been defined to include advanced analytic techniques as well as the actual big data (Russom, 2011). Common software packages, such as SAS Enterprise Miner and IBM SPSS Modeler, allow more data scientists access to the power of big data analytics by simplifying the analytics.

Organizations are seeking employees with greater analytical and technical skills than previous generations. The modern employee, many of whom are digital natives, or those born into a world with ubiquitous computing, have a changing view of technology (Prensky, 2001). The connected nature of digital natives both fuels the amount of data available for a business, but also changes how strategic decisions are made for an organization.

The combination of big data and digital natives in the workforce presents an interesting paradox for businesses. There is a new abundance of information to be gained and a new reliance on technology for decision making. In the short-term, these seem to align, however in the long-term, there will be a side effect of technology dependence that makes decision makers unable to use intuition or common sense. Cognitive effort theories indicate that a reliance on previous technologies results in reduced effort instead of increased quality of outcomes on the part of the employee (Todd & Benbasat, 1992).
Through the lens of General Systems Theory (GST), this research addresses a gap in the literature concerning the impacts of big data and technology dependence on long-term business strategy. Generally, GST seeks to address problems from a holistic perspective to better understand the relationships between the parts of a system (von Bertalanffy, 1972). More specifically, causal maps allow representation of these relationships in a meaningful way that help draw attention to discrepancies that are not intuitive. Combining these tools with a cognitive effort perspective and the theory of technology dominance, this research proposes that big data use follows the “shifting the burden” archetype to show that big data analytics might not always be the best solution to decision making.

This research outlines a clear concern for companies as we move forward with big data analytics: what is the appropriate balance of big data analytics strategies? Organizations are encouraged to pursue big data analytics, even if they are not yet capable of handling the behemoth that is big data. Current research in the area of big data analytics simplifies the contribution of big data, which this research expands.

LITERATURE REVIEW

An Overview of General Systems Theory
The founding of GST is historically credited to Ludwig von Bertalanffy whose academic pursuits were in biology (von Bertalanffy, 1972). In his work, he cites the ideas to ancient times when Aristotle propositioned that a whole could be greater than the sum of its parts (von Bertalanffy, 1972). Despite his biological sciences background, a key focus of GST is the interdisciplinary nature of all problems (Boulding, 1956). In practice, GST research not only spans several disciplines, it attempts to bridge the differences into a common language and develop models of behavior and structure that can be used to explain phenomena on a macro level (Wilby, 2006).

One available output of GST is a causal map. A causal map seeks to show relationships between variables in a given context (Burgess, Clark, Hauser, & Zmud, 1992). Causal maps are an ideal tool for explaining complex organizational behaviors to participants in an organization that might not be able to overcome their own innate human biases (Burgess et al., 1992). While very useful, they do lack some detail and explanation of how the system behaves over time (Lane, 2008). Often, users that are unfamiliar with causal maps will find them difficult to understand, so to provide practical value it is good to use a cycllical process and several participants to help develop and explain the maps iteratively (Burgess et al., 1992).

Some researchers use causal maps as a starting place for understanding a complex situation, and then later develop them into stock and flow diagrams (Homer & Oliva, 2001). Stock and flow diagrams depict the same concepts as causal maps, however they clearly indicate where variables accumulate over time and how the detailed dynamics of the system evolve (Lane, 2008). These diagrams can be created with complex mathematical equations underlying the relationships and then simulated to show how the system may behave (Forrester, 2003). Because research on the intricacies of the impact of big data on organizational behaviors has not been heavily developed, a causal map is used to fill the gap in current research, with the potential for later development of stock and flow diagrams.

Big Data
The problem of big data is not new; however, the extent to which this data is captured and stored is increasing constantly, thanks to ubiquitous computing. As previously stated, big data is not just volume but includes the 3Vs (Gartner Inc., 2013). While the 3Vs describe what big data is, big data is not inherently useful based on this definition. Value is sometimes considered a fourth “V” of big data, because it is an essential part of why big data should be used by
organizations (Watson, 2014). For big data to be useful, analytics must be performed to discern information from the data.

The insights possible from big data are seemingly endless if the proper tools and people are able to translate the data into information. Technology is the only solution for handling big data, as it is too massive to be valuable without the assistance of computers. Technology has a legacy of making complicated situations manageable. As an example, Decision Support Systems (DSSs) were developed in the 1970s as tools to help users make better decisions (Watson & Wixom, 2007). Organizational decisions are often complicated with many different factors to consider at once, and big data is a prime example of aggregating many variables. Paradoxically, the technology that allows for big data might allow for both insights as well as overload (Clark, Jones, & Armstrong, 2007). Analytical decision making involves using analytics instead of intuitions to make decisions (Davenport, Harris, & Morison, 2010). Usually these decisions are judged based on the quality of the decision. Decision quality is not objective and can be the outcomes of the decision, the satisfaction of the decision maker (Lilien, Rangaswamy, Van Bruggen, & Starke, 2004), or the efficiency of the decision (Speier & Morris, 2003).

In the 1990s, DSS focus turned towards Business Intelligence (BI) (Watson, 2014). DSSs have similar traits to BI, and BI is sometimes considered a subfield of DSSs (Kowalczyk, Buxmann, & Besier, 2013). BI is a term that describes a large set of tools used for statistical analysis, data storage techniques, and modeling techniques (Seddon, Constantinidis, & Dod, 2012). BI is important in understanding big data because often BI uses big data to draw insights for the business environment.

The term big data was solidified in BI publications around the year 2011 (Wixom et al., 2014). Some argue that big data analytics is the second half of the BI process; while BI is getting data in, big data analytics is getting information out (Watson, 2014). Others see the two terms as interchangeable. While the two are certainly related, big data presents separate problems from general BI. One of the most significant concerns with big data is the inability to hire enough people that can handle the 3Vs properly (Wixom et al., 2014). The amount of classes and programs to educate students about BI and big data tools are increasing, but practitioners still report a shortage of knowledgeable people to provide business value from big data (Wixom et al., 2014).

A Reliant Workforce

Big data continues to grow in importance, and simultaneously the workforce is fundamentally changing. Many students today, and therefore future employees, are termed digital natives. This generation of people were raised completely engrossed in the use of technology (Prensky, 2001). In contrast, people who might be just as proficient at using technology, but had to learn how to use it manually are digital immigrants (Prensky, 2001). Brains of digital natives develop differently than digital immigrants (Prensky, 2001). Arguably, digital natives see no need to memorize any basic information, because a person can always look up any information on a nearby smartphone (Pogue, 2013). Because of the changing workforce, business operations are evolving because these users are comfortable with relying on modern technology and they make up the bulk of the upcoming workforce.

The use of technology can be a blessing and a burden. Technology can improve decision quality, in fact decision aids should be used to assist professionals with complex tasks (Arnold, Collier, Leech, & Sutton, 2004). For instance, DSSs improve decision quality compared to using unsophisticated tools (Lilien et al., 2004). Others argue that for the technology to truly be effective, it should reduce effort and increase accuracy of decision makers (Todd & Benbasat, 1992). If it does not meet one of these two goals, it will not create value for the company.
Contrasting, technology might also decrease the drive of the user, particularly if the user does not understand how the technology works. Prior research on cognitive effort theory indicates that technology is often designed to reduce cognitive effort, which is desirable to the brain. Despite technology assistance, analysis of big data is complex because of the 3Vs. The decision situation is constantly changing and growing. The cognitive effort perspective, first argued by Payne (1982), claims there is a cost and benefit to every cognitive task. Just as with purchases, thinkers must make a tradeoff between effort and quality of the outcome (Payne, 1982). Using big data analytical technology, a user should be able to reduce cognitive load, which leaves energy to be absorbed somewhere: either reinvested into more analysis, possibly with more variables, or saved by the analyst. DSS research shows people do not reinvest effort saved by the benefits of technology, but just finding a satisfactory answer instead (Todd & Benbasat, 1992). With search technologies, research shows, those who are given a query as a starting search line resulted in more errors, less time, and more confidence in the query than those without a starting point (Allen & Parsons, 2006). Put simply, those who put forth less effort did not have as good of results as those who started from nothing. By not reapplying the effort to handle more complex problems, the employee is in effect depending on the technology for future solutions as well.

The dependence that is created through a lack of reinvested effort can be partially explained by the theory of technology dominance (TTD). TTD recognizes an incongruence between a user's skillset and the complexity of tasks which impacts reliance on technology (Arnold & Sutton, 1998). The key ideas in this theory center on experience with a task and a decision aid, complexity of the task, and the cognitive fit between the user and the decision aid (Arnold & Sutton, 1998). If a person begins to form a reliance on a decision aid, this will cause deskilling (Arnold & Sutton, 1998). Deskilling is a term to describe a growing dependence on technology that results in the decision maker losing understanding of the decision process, which ultimately results in an inability to make a quality decision in the future (Mascha & Smedley, 2007).

TTD has been researched in many decision aid contexts, but not in regards to big data analytics specifically. Research indicates that decision aids might negatively impact inexperienced decision makers (Arnold et al., 2004). Some researchers find that providing feedback on the decision aids functionality can mitigate some of the concern with deskilling (Mascha & Smedley, 2007). Contrasting, some find that experience does not impact decision quality, however inexperienced users are more reliant on the decision aid (Jensen, Lowry, Burgoon, & Nunamaker, 2010). These studies indicate using technology might debilitate a user from making a quality decision without the aid.

The effect of technology dependence is a heavily debated topic in Information Systems research, and often a poorly defined concept. Some use the term technology dependence, while others use addiction, excessive use, problematic use, and some stop at habit; the consensus is only that there is not one definition (Vaghefi & Lapointe, 2014). Despite a debate in terminology, there is generally an idea that it is a problem if overuse of technology cripples a user's ability to function without the technology. For this research, this concept is the definition for technology dependence.

A logical assumption is, if there are more inputs, and better processes, outcomes will improve for a given situation. Despite the increase of the data (input), and the increasing sophistication of analytical tools (process), there is still little research into if and how users are gaining benefits (outcomes) from big data analytics.

A CAUSAL MODEL OF DECISION MAKING WITH BIG DATA ANALYTICS
Peter Senge (1990) is often credited with promoting causal maps developed out of archetype structures. An archetype is a pattern of system structure that occurs repeatedly across different situations (Wolstenholme, 2003). Complex systems result in poor mental models, often because of counterintuitive behavior. By using an archetype as a frame of reference to begin to shape the situation, the system structure may become clearer. To better understand this structure, the “shifting the burden” archetype is applied to using big data analytics to improve decision making (see Figure 1).

Figure 1: Shifting the Burden Archetype

The “shifting the burden” archetype, put simply, proposes there is a problem for which a solution is developed to treat the symptoms, which creates a side effect and ultimately has a negative impact on the fundamental solution to the problem (Senge, 1990). In this context, the problem is poor decision quality and the symptom solution is analytical decision making. This could be a reasonable solution, although it creates a side effect that makes it harder to resort back to the original long-term treatment that should have been applied originally (Senge, 1990). In this case, relying only on analytical decision making creates a side effect of big data dependence, which makes it harder to ultimately make independent decisions. Through this archetype, I will discuss the role of big data in organizations (see Figure 2).

Figure 2: Big Data Analytics Archetype
To better understand this archetype from a GST perspective, individual loops are explained in detail. First, “shifting the burden” revolves around a problem behavior: the need for a higher decision quality (see Figure 3). Organizations might have a poor decision quality, or might have a poor *perceived* decision quality because competitors are using cutting edge decision tools. Currently, there is a significant push for more analytical decision making to improve decision quality (Davenport, 2006). When applied appropriately, analytical decision making can be a huge benefit resulting in better decisions (Davenport, 2006). Big data innately requires technology, however big data analytics is treated as somewhat of a golden ticket to problem solving. Using big data analytics allows for new insights that otherwise might not be intuitive for a decision maker. The benefits of big data analytics are obvious and most companies see big data analytics as an opportunity for the future (Russom, 2011). As decision quality of the firm increases, the organization will begin to use more analytical decision making. Discovering these new insights increases the perceived value of big data analytics, and thus the organization will want to do more big data analytics.
The downside of analytical decision making is that not all decisions should be made using big data, such as non-routine decisions (Kowalczyk & Buxmann, 2014). Data analytics relies on past occurrences, and as a result, if a situation is an anomaly, it would not be wise to try to use past data analysis to arrive at a solution. Analytics provide sound reasons for making decisions, but not all companies are able to continuously gain benefits from big data analytics given the current state of their business. Presumably, more analytical decisions should positively impact the decision quality, however this is not always the case. Part of decision quality includes the financial cost of the decision and the time it takes to make the decision. The use of big data analytics for simple decision tasks would also hurt decision quality if the decision could have been reached by cheaper, faster, and simpler means. Instead, big data should be used in appropriate decision tasks that relevant data is available to determine factors that are important for a problem that has happened before (Kowalczyk & Buxmann, 2014). Another primary concern with big data comes from making decisions using dirty data (Kopytoff, 2014). If the data is not cleaned properly, there is a high potential for misleading information, and data cleaning can be time consuming and costly (Li & Joshi, 2012). Making decisions with analytics when it is not appropriate can have a negative impact on decision quality. As a result, this creates a balancing loop as pictured in Figure 3.

Analytical decision making can certainly be used only in appropriate situations, however, even in this case it creates a side effect (see Figure 4). The side effect is a dependence on big data. Companies are facing a shortage of people able to perform the required analytical tasks, and as a result, they are trying to replace the people with automated statistical tools (Fitzgerald, 2014). Because there is a shortage of people who can effectively make analytical decisions, the tools might be used by inexperienced decision makers. This idea is addressed in the TTD and suggests that this will create a reliance on the tool (Arnold et al., 2004).

For the organizations that have high quality data scientists who can perform appropriate analytics, there is still a concern. As long as the data scientist constantly reinvests the effort that
is saved from using the technology back into using more data or making more complex decisions, then the result will be good. Unfortunately, as prior research suggests, saved effort from technology use is often not reinvested to improve decision quality (Todd & Benbasat, 1992). Instead, the user just becomes more dependent on the technology.

People are relying on technology at the individual level as a cognitive crutch that no longer requires effort (Pogue, 2013). When people rely on technology, they no longer exercise their brains and lose some ability to process complex problems, or deskill (Mascha & Smedley, 2007). At the organizational level, some companies are relying heavily on analytics to make decisions (Davenport et al., 2010). Using System Dynamics, deskilling was effectively shown as a concern in IT outsourcing in prior research (McCray & Clark, 1999). Instead of outsourcing to another company, big data analytics are being “outsourced” to automated tools that can also hurt the user’s knowledge (Fitzgerald, 2014). While big data can improve decision making, there is still a need to have quality decision makers that know how to make a quality decision with the aid of data (McAfee & Brynjolfsson, 2012). A shortage of people, as well as a lack of understanding and effort by current decision makers, creates a dependence on big data that further creates a problem for a company.

The dependence on big data results in a negated ability to make decisions appropriately. If people are reliant on the big data analytics and not reinvesting effort, or if there are not proper data scientists performing the analysis, the ability to make decisions without analytics will suffer. As the company grows in dependence, big data analytics will become the default option, even when it is not the proper solution to a decision-making task. The more dependent a company is on big data analytics, the harder it will be to correct the behavior of misusing big data analytics. This problem continues to perpetuate as a greater need for increased decision quality leads to more analytical decision making, which in turn creates greater dependence on big data.

The solution to the problem is to apply the appropriate decision making techniques with the appropriate amount of resources that fit the task (see Figure 5). Put simply, hiring data scientists that can properly reinvest effort to increase decision quality. The primary reason this is not the chosen solution for some organizations is because it can be a delayed effect. For instance, hiring sophisticated data scientists that can properly turn big data into a valuable resource might take significant time, money, and training.

Using big data analytics will likely appear to improve perceived decision quality in the beginning, even if it is misapplied or handled by unskilled statisticians. The belief that the decision is better will send the organization spiraling into the loop pictured in figure 3. However, hiring the appropriate data scientists will more accurately provide a long-term solution. If an organization is experiencing a high decision quality at the moment, the relationship to hiring high

![Figure 5: The Solution](image-url)
quality data scientists will be negative. There is no need to hire data scientists if there is a perceived high decision quality.

Overall, data scientists improve decision making ability in the long-term. Because these employees are sophisticated at handling big data analytics, they are more likely to use the conserved effort and reapply it into decision-making. Further, only using big data for tasks that big data can supply answers to will reduce dependence on big data. Using the most appropriate decision making task would reduce the big data dependence, and instead only encourage it when it is an appropriate tool for a decision. This creates a balancing loop that keeps decision making and decision quality in an optimal balance.

DISCUSSION

The causal map presented in this research explains how big data analytics might not always be the best solution, and instead it might create a dependence on big data technology that is not necessary for simple decisions. The implications of these ideas are two-fold. First, practitioners should take caution in dependence on big data to make decisions. As illustrated, this can lead to a vicious cycle of dependence that begins to inhibit quality decision making. Instead, big data should be used in tasks that are appropriate and some decisions should still be made on intuition and available information.

Secondly, research can benefit from this analysis. This work suggests there are boundary conditions for when big data analytics is the optimal solution. Research should further investigate what are the prime situations for big data analytics, and what simply further perpetuates a dependence that is unnecessary. This should be analyzed in two ways; first a simulation model might further indicate counterruitive behaviors of the behaviors presented that might result in bigger problems that it is hard to foresee using only mental models (Sterman, 2001). Secondly, reductionist research might provide other insights into these relationships. Using other methodologies, such as empirical analysis of practitioner data, would be a useful way to validate these ideas. Further research into the relationship between big data and decision-making could shed new light on the use of technology as a decision aid.

Specifically a few research questions might include:

- What are the limits of big data analytics?
- In what contexts is big data analytics most efficient at creating value for an organization?
- How can organizations encourage reapplication of conserved effort in big data contexts?

By exploring some of these questions, new ideas will likely arise that will make big data analytics more valuable to companies, and thus make an impact on determining what is the best solution for an individual context.

LIMITATIONS

All systems theory models are subject to boundary conditions that can be just as important as the causal map itself (Wolstenholme, 2004). The model presented here is no exception and was designed with boundary conditions in mind. First, this work assumes that there are not an adequate number of data scientists available and that there is a delay before future scientists can be integrated into an organization. Second, not all individuals behave in the same way, and some employees will reinvest effort appropriately instead of conserving. This would make the dependence on big data less of a concern if the energy is reinvested appropriately. Third, this model is bounded by the investment capabilities of the organization. Big data analytics is not applicable for every organization, as it can be costly and too sophisticated for some smaller organizations. Finally, decision-making is a complex task in itself, and as a result there are
several outside factors that go beyond the scope of this research that might play a significant role in big data dependence.

Further, without simulation, causal maps are arguably untested. Many systems theorists believe that the true nature of the system can only be understood with simulation, and thus a causal map would not accurately portray the complex dynamics (Homer & Oliva, 2001). Despite the concern, causal maps are a useful place to begin to understand complex situations, particularly with behaviors that are not easily measured, such as decision making (Coyle, 2001). As a result, the causal map presented here provides a good background structure for understanding the problem and furthering the ideas in this area. More traditional empirical research in this area, with the consideration of the issues brought out by the causal map presented, would likely provide valuable insights to business practitioners.

CONCLUSION

The causal map presented in this research attempts to explain the role big data can play in an organization and the counterintuitive behavior it might cause. Primarily, the nature of big data might lend itself to creating a dependence that can ultimately have a negative impact on decision making within an organization. This work fills a gap in big data research by considering the limits to the value of big data analytics and considering inappropriate contexts to use big data. While the opportunities big data presents can be limitless, like all good tools it should be used within the contexts that it is most suited for.

REFERENCES


Gerhart A systems theory perspective of data analytics implementation


Gerhart

A systems theory perspective of data analytics implementation


