

Factors Impacting the Influence of Analytic Capabilities on Organizational Performance in Higher Education

Cory A. Campbell
Indiana State University
cory.campbell@indstate.edu

Philip A. Cola
Case Western Reserve University
philip.col@case.edu

Kalle Lyytinen
Case Western Reserve University
kalle.lyytinen@case.edu

Abstract

In response to changing fiscal needs and opportunities, higher education institutions have adopted new ways to use financial information for improved decision making. Drawing upon resource based theory we examine the connection between university level data analytic capabilities and organizational performance. We posit this relationship to exist through a serially mediated path of data-driven culture and data quality. The study provides empirical evidence that establishing a data-driven culture contributes to data quality which together result in increased organizational performance. The serial mediation pathway creates a positive effect between data analytic capabilities on organizational performance. This is critical information relative to both resource based theories and practical implications for higher education relative to beginning the investment cycle at the organizational culture level related to use of data.

1. Introduction

Higher education has an ever growing and significant impact on society. In addition to cultivating future leaders and promoting societal advances such as tolerance, evidence-based politics, civic discourse enabled through academic research and teaching, higher education possesses a significant economic footprint. In the United States, higher education impacts both local and national economies with revenues and expenses eclipsing \$500 billion per year [1].

At the same time, economic pressures are heightened more than ever in higher education institutions as their funding sources have significantly tightened. More than 50% of the states in the U.S. have adopted some format of performance-based funding where financial support is tied in some type of student success outcome [2]. Public institutions also compete for reduced state appropriations due to increasingly constrained state budgets and lack of political support to garner greater economic resources. At the same time for private institutions, ‘net’ tuition has been shrinking as discount rates have soared over the last decade. According to a national survey, the discount rate for

first-time, full-time college freshman at private institutions has increased from 39% to 52.2% [3] meaning that internal institutional aid covers now 52 cents for every dollar of tuition for private institutions. For both public and private institutions, the disruptive and dynamic shift in funding sources has pushed universities in an uncharted territory. Sustaining current levels of revenue will increasingly challenge universities in the future as demographic trends show that the number of high school graduates will start decline through the 2030 academic year [4].

Health and welfare concerns, funding constraints, declining enrollments and unfavorable demographic trends will serve as a call to action to ensure that schools adapt to the “new” normal. Institutions that have historically made incremental changes to their fiscal policies will require a paradigm shift in how they organize and manage the universities. These organizations need to adopt new sets of metrics for evaluation and must respond swiftly and flexibly. In this new environment, data quality and extensive use of data analytics has been viewed as one ‘silver bullet’ whereby schools can become more efficient and learn to better manage budgets and resource allocation.

Through data analytics, institutions are expected to be able to recognize trends, ask more salient “what-if” questions, discern novel correlations, apply predictive models, and use financial models to support new initiatives and programs [5]. However, higher education has historically struggled with strategic information technology (IT) alignment. Overall, there is little prior research on effects of data analytics on higher education performance that combine both financial and non-financial measures [5].

In the past, IS research has leaned strongly on Resource Based Theory (RBT) to explain the benefits of deploying IT based competencies to generate value for the organization [6-8]. In this study, we examine the conjecture that resources which enable higher education institutions to improve their performance include data analytics related resources beyond such investments in data. We posit that these resources are valuable, imperfectly imitable, non-substitutable and occasionally rare and cover the quality of data, organizational culture, and forms of deploying analytics capabilities. Therefore, our research question

is: *To what extent does analytics capabilities drive organizational performance in higher education?*

1. Literature Review

Higher education institutions have for some time recognized the need for improved financial effectiveness. Due to multiple fiscal pressures, the current economic climate places a premium on making effective decision making and increased financial transparency. The National Association of Collegiate and University Business Officers (NACUBO), stresses that high performing colleges and universities are increasingly governed by sophisticated business administrators who equip the organizations with improved financial planning capabilities and actionable information [9]. Despite the fact that institutions have recognized the need to use data analytics extensively, little progress has been made towards this goal [10] while they have implemented ERP systems to leverage activity efficiencies and eliminate data redundancies. Currently, these systems are not being used to the greatest possible extent for organizational benefit given their improved data quality and widened scope of utilization [5, 10].

Data analytics is a resource (ultimately resources), that are valuable, rare or unique, inimitable, and organizationally embedded (VRIO) [11]. This is especially true in higher education where attracting and retaining degree earning students is the lifeblood for any institution of higher learning. A joint statement of national associations that support institutional research, information technology, and business officers highlights that “data are an institutional strategic asset and should be used as such”, suggesting this is a valuable resource that is rare [12]. The joint statement goes on to suggest strategies of making an institutional commitment and that analytics is a ‘team sport’, confirming from earlier reports that this is both inimitable and organizationally embedded. The primary premise of RBT theory is that the possession of certain value generating resources shapes the effectiveness and use of the organization’s resources. Thus, RBT provides a lens to understand the effects of the use of data analytics and related resources in conjunction with other organizational resources [13]. In this perspective, RBT concepts connect IT value to encompass multiple value creating aspects. Over time, firms also develop IT based VRIO capabilities through a series of linked strategic decisions focused on information technology investments that integrate IT into organizational processes and knowledge [14].

IT based resources are “valuable” as the changing environment places a premium on information, which also drives organizational efficiency. In the wake of

unfavorable demographic and fiscal constraints which many institutions of higher learning now face, the chief business officer of a large statewide systems describes how placing premium on “(data) analysis could help figure out how we can get more efficiency out of the operation” [15]. Resources are “rare or unique” as demonstrated by the limitations to which processing capabilities are constrained by a human capital problem. As one CIO in the university described, “I am looking for a needle in the haystack because there is not many people who can bridge those two (accounting and technology) worlds” [15]. Higher education is also limited vis-a-vis human capital as those working in roles as report writers and analysts tend to be home-grown [15]. Resources are costly to “imitate” as there is no one size fits all solution but rather universities need different types of process improvement for their data analytics [16]. Therefore, the factors of valuable and rare or unique are necessary, but not sufficient unless there is a component of inimitability.

This is a challenge in the wake of the big data era as managers seemingly need to separate relevant from irrelevant information for their decision making [17]. Organizations must rely on their agility which defines an organization’s ability to adopt or adapt business processes to achieve speed, accuracy, and cost economy [14]. Resources are also “organizationally embedded” insofar that culture and learning capacity become deeply embedded in the organization, which has also been confirmed in the higher education [15]. The need for the higher education sector to quickly adapt to the changing environment highlights the important role of utilizing related institutional capabilities in the form of digital options and agility [14]. In this regard data analytics conforms conceptually to RBT ideas as its VRIO principles of resource deployment are applicable.

The promise of data analytics is compounded by the increased volume, variety, and velocity of data. In this regard the two technical limitations of big data analytics are storage and processing capacity. Not only does business analytics positively enhance information process capability, but also having a simultaneous data-driven environment forms an antecedent of using business analytics effectively [18]. Grover et al [19] suggest that pairing structured and unstructured data together can yield insights, which have never been considered before. Since there are differences across industries regarding the use of business analytics applications [18], exploring the possibilities of data analytics is important for higher education.

The foundation behind RBT is that efficiency differences explain why organizations vary and that distribution of resources and capabilities is a key

differentiator among competing organizations [20]. RBT posits that organizations are able to build sustainable competitive advantage which yields superior returns by assembling both physical and non-physical assets as resources with the capabilities and processes to add value [8]. Others argued that capabilities are not a resource because capabilities are not observable and therefore intangible, capabilities cannot be valued, but rather are dependent on wherever the capability is housed within the organization [21]. Gupta & George [11] use the tenets of RBT to demonstrate how big data analytics are composed of tangible, intangible, and human capabilities. Other researchers have leveraged RBT to examine how firm performance is a product of IS resources and capabilities [22]. Bhatt & Grover [23] present a model for IT capabilities based on the premise of uniqueness that can help build competitive advantage. RBT focuses on efficiency-based differences to provide value and the primary mechanism for an organization is to maximize resources. At the same time RBT offers a lens through which one can consider the role of information systems to create a sustainable competitive advantage [8].

IS studies show that resource utilization is relevant to identify the true value of analytics capabilities [11, 19, 22]. RBT has been used to explain how data culture forms an antecedent to knowledge management. This suggests that as a resource, culture may be either a catalyst or a hindrance [24].

Therefore, the Research Model in Figure 1 includes data quality, data-driven culture and organizational performance as endogenous variables. Kwon et al [25] suggest that greater efficiency of managing data quality can lead to other improved firm capabilities as data quality is an organizationally embedded resource. Thus, it seems logical to expect the transfer of effects from analytic capabilities to organizational performance would at least in part be impacted by data quality. Furthermore, Grover et al (2018) suggest that data-driven culture can provide predictions about where a company is going, which would also indicate that data-driven culture would impact the effects of analytic capabilities on organizational performance.

Organizational performance generally is measured by assessments of effectiveness, efficiency, productivity, quality or even innovation [26]. In this study, we adapt these ideas toward how well institutions of higher learning recruit, retain and ultimately graduate students. In order to summarize these important institutional performance outcome measures, we define organizational performance as being equal to net revenue per student. Past research suggests that criterion for the dependent variable,

organizational performance in this study, in a resource based framework will need to convey performance assessments, competitive assessments, and conduct such assessments over time [8]. Internal and external assessments of the institutions of higher education must be performed more frequently using data analytics tools in order to better understand if data capabilities help to change the culture, data quality, and overall performance. Organizational performance in the higher education sector fit this definition and paradigm of more frequent assessment on investment, quality and culture. Furthermore, the organizational level as the unit of analysis would also be appropriate for the use of RBT. Thus, in this case we look at data from an aggregated organizational level to try to better understand the impact of analytic capabilities on organizational performance as mediated by data-driven culture and data quality. These are important factors to understand in the contemporary business modeling for institutions of higher education.

Figure 1: Research Model



Hypotheses Development. The literature review sets the theoretical framework for the Research Model presented in Figure 1 above and below the causal logic for the hypotheses tested in this model is provided and supported by additional literature streams emanating from information systems in higher education.

Direct Effects. The dependent variable, organizational performance of higher education institutions, is posited as the key performance indicator for this study and is defined as the net revenue per student for each university that has data included herein. In the research model, this construct provides a measure of organizational performance of the university while recognizing for the fact that public and private schools have different components of revenues and expenses.

Analytics capabilities relates to the investments, projects, and activities of enterprise IT departments [27]. Analytics investments have been examined by IT spending (amount), IT Strategy (type of investment) and IT strategy/capability (asset management) [28]. Analytics capabilities are a resource and as such, it creates capabilities and value for the organization and the use of analytics requires investment in people and technology. Conceptually we define data analytics capabilities as the use of data, statistical analysis, and

explanatory and predictive models to gain insights and act on complex issues.

Several studies show a direct link between analytic capabilities and a performance outcome measure [11, 22, 23, 29, 30]. Likewise, this study considers the effects of analytics capabilities in higher education, another underrepresented sector in analytics research. Analytics capabilities creates value, thus we expect to find a direct relationship between analytics capabilities and organizational performance.

Hypothesis 1. *Analytics Capabilities is positively associated with Organizational Performance.*

Mediating Effects. Data quality refers to the consistency and comprehensiveness of the data and arises from its sources, which may be internal or external to the organization. Data quality is the output of these various financial and non-financial inputs that are aggregated together for access and analysis. For this study, data quality will be defined as accurate, timely, complete, and consistent data [31]. The amount of resources, such as analytics capabilities and data-driven culture, influence data quality. Thus data quality is both valuable and can be costly to imitate. The effectiveness of the data quality is contingent on how the data is employed, which makes data quality rare.

Firms gain competitive advantage through their internal and external strategies. Prior RBT research suggests that the evolution to data analytics depends on the quality of data within the organization [25]. To sum up, data quality is valuable, non-imitable, rare, and organizationally embedded and can be leveraged as a resource to improve analytics capabilities outcomes. However analytics capabilities remain imperfectly mobile insofar that it is difficult to acquire in resource markets or to develop internally [23]. Using RBT, we expect that data quality forms a critical element contributing to the transfer of effects from analytics capabilities on organizational performance. Therefore, we hypothesize:

Hypothesis 2. *Data Quality mediates the positive relationship between Analytics Capabilities and Organizational Performance.*

In the context of this study, data-driven culture is defined as patterns of behaviors and practices by organizational members who share beliefs that having, understanding, and using certain kinds of data and information plays a critical role in the success of their organization especially related to decision making practices that impact organizational performance [32]. Therefore, data-driven culture is valuable as well as a non-imitable resource.

Prior research has suggested that matching the culture inherent to the organization and the assumptions embedded into the information systems is

a critical success factor [33, 34], which could be problematic for what some have depicted as a laggard sector with the adoption of analytics [5, 10]. Elbashir et al [35] show that the relationship between top management support is “crucial” to determine the sufficiency of resources such as time, finance, information, and human resources. Prior research on analytics using an RBT lens finds data-driven culture to be a strong moderator to create value for innovative companies [19]. It may be data-driven culture that is the missing element of design, innovation, and creativity within a domain such as higher education where tradition and culture are deeply embedded in the ability, or inability, to organize time and space. This element of culture has not only served as a predictor of organizational performance, but is also postulated as a potential mediator of the transfer or effects on organizational performance. Thus, we hypothesize:

Hypothesis 3: *Data-Driven Culture mediates the positive relationship between Analytics Capabilities and the Organizational Performance.*

While it is plausible the data-driven culture and data quality may exist independent of one another, it remains unclear as to whether data-driven culture and data quality mediate separately yet in tandem or whether the sequence of the factors matters in a serial mediation model. Using the RBT lens to consider data-driven culture as well as data quality, the sequence of constructs would enhance the value of viewing them through the VRIO framework. Therefore, through this lens we speculate about the plausibility that the constructs of data-driven culture and data quality form a causal chain (i.e., Analytics Capabilities→Data-Driven-Culture→Data Quality→ Organizational Performance, in a model referred to as serial mediation [36]. Serial mediation is important to explore because the differential impact on organizational performance seems to matter when thinking about when and how to begin analytics capabilities initiatives [37]. We postulate that the establishment of culture in an organization towards the collection and utilization of data for decision making precedes the quality of the data, which in turn collectively impacts organizational performance. Therefore, we posit the following:

Hypothesis 4: *The relationship between Analytics Capabilities and Organizational Performance is serially mediated by data-driven culture and data quality.*

Controls. Retaining students and successful matriculation of students are core mission values for higher education. Two widely accepted throughput metrics include graduation rate [38] and the student retention rate [39]. These two control factors are standardized through required reporting to the

Integrative Postsecondary Education Data System known as IPEDS. An additional control variable for the size of the institution is included in the model. The size of the institution must be considered as capabilities differ greatly by sheer volume of activity with an organization.

While both public and private institutions face financial stress, public institutions have seen operating revenue increasing in line with expenditures, whereas private institutions face significant gaps as increases in revenue are not keeping pace with rising costs [1]. From the 2009-10 academic year to the 2014-15 academic year revenues increased 22% for public institutions, while expenditures grew by 19%. However, their private institution counterparts saw only 6% increase in revenue, but expenditures ballooned by 24% during that same timeframe [1]. This amplifies the importance of data quality in private institutions facing unsustainable trends in net operating revenues. Disproportionate increases of expenditures over revenues in private schools would warrant exploration of group differences between private and public institutions. Therefore, the research design controls for private versus publically funded institutions.

2. Research Method

Design and Measures. Co-variance based structural equation modeling (SEM) utilizing SPSS and AMOS (v26) was employed to test the hypotheses. To test the serial mediation hypotheses, the direct, indirect, and total effects of data-driven culture and data quality on organizational performance were estimated. Integrating two models of mediation through data-driven culture and data quality yields a three-path meditation model [36]. To consider the significance of the indirect path, regression analysis with a bootstrapping technique is used [36]. The advantage of this approach allows for isolation of each mediator's indirect effect as well as the indirect effects in a sequential pathway from the independent variable to the dependent variable in a series [36].

Due to the fact that institutions of higher education accept student loans and other sources of federal funding, financial and non-financial data is publically available for both public and private institutions. Under U.S. federal law, any school, which received Title IV funding (e.g., student loans), is required to report key information including financial and enrollment information.

The study measures of Analytics Capabilities, Data Quality, and Data-Driven Culture originate from survey data collected by a nonprofit organization focused on the role of technology in higher education. This survey data is combined with publically available data from the Integrative Postsecondary Education Data System (IPEDS) data.

We utilized IPEDS data the computation of the net revenue per student, a proxy for Organizational Performance. The difference between core operating revenue and core operating expense¹ is be divided by the student full-time equivalent² to compute net revenue per student. Absent of direct profit metrics in the higher education, net revenue per student will serve as a reasonable proxy for organizational performance. The National Center for Education Statistics recognizes that changes over time in accounting definitions have made it difficult to compare private and public institutions [40]. Growing interest in higher education finances, including different ways that institutions can do more to promote cost effectiveness, has led to improved reporting for revenue and expenditures [1].

Net revenue per student was derived using IPEDS data. This metric is relevant because non-profit colleges and universities have begun to act more like organizations in the for-profit sector [41]. IPEDS data would ensure uniformity, including operating revenue and expenses as well as enrollment.

Despite being key performance indicators on various education statistical websites, the higher education academic literature remains silent on net revenue per student measures at the organizational level. The current literature has sought to inform policy and suggest specific student level outcomes [42] rather than considering financial performance measures. For the context of this study, the net revenue per student is the difference between core operating revenue and core operating expense divided by the student full-time equivalent. Thus, the calculation for net revenue per student is calculated per the following formula:

$$\text{Net Revenue Per Student} = \frac{(\text{Core Revenue}^1 \text{ less Core Expense}^1)}{\text{Full Time Equivalent}^2}$$

Data Collection and Sample. The research framework included the Data Maturity Index from Educause. Over 1,900 colleges and universities, 350 corporations, and numerous state and federal agencies are members of Educause. The study sample consists of 466 public

¹ Core Revenues and Core Expenses as defined by IPEDS include GASB standard 34 for public institutions and FASB standards for private institutions.

² Full-time Equivalents (FTE) as defined by IPEDS is a single value providing a meaningful combination of full-time and part-time students.

and private institutions which completed the 2014 survey. The final data set contained twelve items including categorical moderating and control variables. Analytics capabilities, data-driven culture, and data quality reflect the 2014 fiscal year, whereas the dependent variables and its control variables are from the 2015 fiscal year. Individuals who completed the surveys are most knowledgeable about the overall state of analytics at their respective institutions, such as the CIO, director of institutional research, or the officer responsible for institutional performance management.

Pairing of Educause survey respondents and IPEDS data allowed for further analysis. Computation of net revenue per student utilizes IPEDS data. IPEDS data would homogenize the data that schools report as core revenues, core expenses, and enrollment. The population includes plus or minus two standard deviations from the organizational performance construct as defined above. There were eight institutions excluded from the sample population because the net revenue per student was greater than plus or minus two standard deviations, leaving the final sample size of 466 institutions.

Data Analyses. Screening of this data occurred to investigate univariate assumption that includes homoscedasticity, skewness, and kurtosis. Following data screening, we conducted both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to ensure the adequacy, reliability, and validity of the data, prior to the utilization of SEM for hypotheses testing.

Univariate Assumptions. During data screening, we examined data for skewness and kurtosis. All variables exhibited some modest homoscedasticity, but fell within acceptable ranges [43]. Data was sufficiently large to reduce the minimal effects of skewness and kurtosis when using Likert scales.

Measurement Model. In order to validate the survey data collected from a secondary source, we conducted an Exploratory Factor Analysis (EFA) to determine if our factors loaded adequately in this model. The extraction method used was Maximum Likelihood with Promax rotation. Factors were eliminated one by one as the EFA output was continuously re-assessed using a threshold for commonalities of 0.40 as per MacCallum et al [44]. Once the commonalities were acceptable, the freely estimated model loaded to three factors (analytics capabilities, data quality, data-driven culture), which are listed in Table 1 including each reliability measure. The Bartlett's Test of Sphericity is significant and the Kaiser-Meyer-Olkin (KMO) is adequate at 0.862. The three factor EFA (analytics capabilities, data quality, and data driven culture) explained 55.9% of the variance, with non-redundant residuals of less than 4.0%. The EFA was done for

reflective constructs only, so the measurement of organizational performance as measured by net revenue per student was not subject to the EFA because this is not a reflective measure.

As evidence of convergent validity, all the loadings in the pattern matrix are greater than 0.50. As evidence of discriminate validity, the questions comprising the constructs have no cross-loadings in the final pattern matrix. As evidenced by the Cronbach Alpha > 0.70 , there is strong reliability for the factors.

The final pattern matrix from the EFA was used to produce a Confirmatory Factor Analysis (CFA) using AMOS. Examination of modification indices revealed several strong relationships between error terms in the model. As such, we removed one item from data-driven culture to improve the model fit. The CFA evaluated a 12-item factor solution. Final model fit statistics were adequate: Cmin/df=2.209, GFI=0.966, CFI =0.976, RMSEA =0.051 and PClose =0.438. The SRMR was 0.0335. With adequate fit established and satisfactory factor loadings we next evaluated model validity and reliability.

Determining the Average Variance Extracted (AVE) of each factor assesses convergent validity in the CFA. The squared correlations between factors are greater than 0.50 [45]; therefore, all are deemed acceptable, suggesting adequate convergent validity. The data also suggest discriminant validity for each of the factors because the square root of average is greater than any inter-factor correlation and the maximum shared variance (MSV) values were not greater than the AVE values. Composite reliability values are sufficient.

Table 1: Correlation Matrix and Validity

	Mean	Std Dev	CR	AVE	MSV	Analytics		
						Capabilities	Data Culture	Data Quality
Analytics Capabilities	3.310	0.185	0.822	0.538	0.345	0.733		
Data Driven Culture	2.570	0.167	0.838	0.634	0.345	0.587	0.796	
Data Quality	2.470	0.039	0.801	0.502	0.250	0.500	0.499	0.709

To test and potentially correct in part for method bias, we included an unmeasured latent factor in the CFA. Adequate model fit with the inclusion of a common latent factor was: Cmin/df=1.631; GFI=0.981, CFI =0.991; TLI=0.977; PClose=0.878, RMSEA = 0.037. The fit was deemed adequate to proceed with testing [46-49]. Thus, a Chi-square difference test was employed to examine whether the non-congeneric model or the null fit the data [50]. An unconstrained model with all loadings from the unmeasured latent factor allowed to freely estimate was compared to models where the latent factors were first constrained to be equal to one another, and then constrained to be equal to zero [50, 51]. The Chi-square difference was 41.64, and degrees of freedom

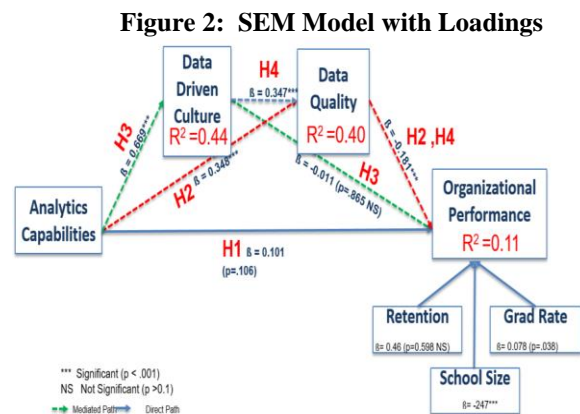
difference was 11, with a $p < 0.001$. Therefore we rejected the null hypothesis and concluded that partially shared variance comes from method bias. The unmeasured latent factor was included in the CFA as well as the structural analysis to account for common method bias.

A curve estimation shows that all the relationships in the model are linear. The variable inflation factor (VIF) values were all less than 3.0, indicating that the variables are distinct and multi-collinearity was negligible [52]. Two responses exhibited an abnormal Cook's distance, therefore these two outliers omitted.

We used AMOS to test the final structural model, which included the unmeasured latent factor where imputed values were included in the structural model. The mediated model with controls was run by using 2,000 bootstrapping resamples using 95% two-tail confidence intervals [36]. Model fit for structural model was adequate: $Cmin/df = 1.376$, $GFI=0.995$, $CFI=0.998$, $TLI=0.958$, $PClose=0.754$. The SRMR is 0.0754 [51,53,54].

3. Results

As shown in Figure 2, percent variance explained for the endogenous variables reveal R^2 values for data quality was 0.40, for data-driven culture 0.45 and for organizational performance 0.11.



Direct Effects. Hypothesis 1 considered the direct effects on analytics capabilities on organizational performance. The direct effect was not significant ($\beta=0.101$, $p=.106$ NS). Therefore, **H1 is Not Supported**. There was no direct relationship between analytics capabilities and organizational performance.

Mediation Effects. Data quality **fully mediates** the relationship between analytics capabilities to organizational performance (H2) as evidenced by the significance of the indirect path ($\beta= -.181$; $p < 0.001$). Thus, **H2 is supported**. The path from data-driven culture to organizational performance (H3) is not

significant ($\beta=0.011$; $p < 0.865$ NS). Thus, **H3 is not supported**.

H4 predicts that data-driven culture and data quality serially mediate the relationship between analytics capabilities and organizational performance. Serial mediation shows a causal chain between the mediators with a specified direction leading up to the dependent variable [36]. We found there to be a sequential pathway in the mediation chain in that a data-driven culture precedes data quality toward improving organizational performance. This is evidenced by the indirect effect of data-driven culture examined here being significant ($\beta=0.347$; $p < 0.001$). Thus, **H4 is a supported** indicating that analytics capabilities requires a data-driven culture or data quality, and moreover in that order, to impact organizational performance.

Control Variables. The core purpose of higher education is to graduate students. However, the rate at which students graduate controls for organizational performance was not significant ($\beta= -0.078$, $p=0.338$). Institutions devote resources to increase retention. Since resources cost money, the expectation is that retention would be an expenditure that would be negatively associated with organizational performance. However, this control variable has no significant effect ($\beta=0.046$, $p=0.598$ NS), whereas size does have an effect on organizational performance. However, this was a negative effect ($\beta= -0.274$, $p < 0.001$). This interesting negative effect with the control of institutional size led us to conduct *ad hoc* multi-group moderation analyses to understand this mechanism of action more fully.

4. Discussion

We set out to better understand the impact of analytics capabilities on organizational performance in higher education. Also, to determine the extent that a data-driven culture and data quality had relative to analytics capabilities on organizational performance for higher education. We found that analytics capabilities did not predict organizational performance in of itself, however, we did find that data-driven culture and data quality both significantly mediated the effects of analytics capabilities on organizational performance. Most importantly though we found that there is a serial mediation pathway indicating the need to establish a data-driven culture first in order for analytics capabilities and data quality to more fully realize the impact on organizational performance.

This has large implications theoretically and practically when there is increased pressure for external funding for both private and public schools, and stakeholder expectations are on the rise. Due to

increased pressure of the expectations, nonprofit organizations are competing with each other for limited funds, nonprofit organizations face increased pressure to be more accountable [41]. In higher education, the external environment is changing the need for further decreased expenditures. The sector faces unfavorable demographic trends before the uncertainty about future enrollments and budget shortfalls as the aftermath of the global pandemic of 2020. The extent to which the impact of social distancing changing the college experience and the effect on student success remains an enigma. Therefore, a critical aspect of the new digital economy is the ability of organizations to realize economic value from a collection of knowledge assets including information and collaboration with other external sources for data to use in decision making [53].

The current study suggests data quality mediates the effect of analytics capabilities and organizational performance. Analytics capabilities alone will not yield the desired results, as there is no direct effect. In order to realize the potential benefits of analytics capabilities, organizations need to focus on improving data quality. One key finding was the negative relationship between data quality and organizational performance suggests that perhaps there is a belief that the quality of data is perceived to be a constraint by those who rely on data to drive decisions. The overall negative response for survey respondents (mean is 2.47 on a 5 point Likert scale) suggests that perhaps there is a belief that the quality of data is perceived to be a constraint by those who rely on data to drive decisions. Higher education is data rich, but information poor. In order to overcome this limitation, focusing on a data-driven culture is the critical success factor. Furthermore, the relationship of analytics capabilities on data quality will yield positive benefits that can be leveraged further through a strong data-driven culture within organizations. In other words, data quality is implicit of data-driven culture insofar that these two must co-exist in order to optimize the overall effect with the data-driven culture being established first to maximize the effects of data quality. This suggests institutions can leverage efficiency of resources and mobilize activity around data quality to achieve better outcomes.

To further assess the strength of the multiple mediation model, a *post-hoc* analysis considered the strength data quality with ($R^2 = .40$) and without ($R^2 = .034$) data-driven culture. Since the questions that comprise data quality are based on human perception, it makes sense that data-driven culture would strengthen the effect on data quality. Note: data-driven culture is an antecedent to data quality because it has no direct effect on organizational performance,

whereas data quality does have a direct effect on the dependent variable.

Data quality alone does have an effect on the dependent variable, but it can be enhanced through an organizational culture that believes analytics will provide better decision making capabilities. Peter Drucker famously coined the phrase *culture eats strategy for breakfast* [54] and this analogy holds true in academia whereby a data-driven culture is proven to be an antecedent to both data quality and organizational performance. The influence of data-driven culture to enhance the effect of data quality on organizational performance has similar effects for both private and public schools. This indicates its importance across the board in the higher education space.

The study contributes much evidence toward the answer to research question. First, this study introduces organizational performance measured by the net revenue per student as a new construct to the academic literature. This study uses net revenue per student as a proxy for an aspect that explains a degree of organizational performance. The strength of this as a dependent variable is the standardization of operating revenues less operating expenses, for both public schools adhering to Governmental Accounting Standards Board (GASB) and private schools following guidance from the Financial Accounting Standards Board (FASB) while controlling for school size as efficiency metric, measured on a per student a basis. Second, this study contributes to the analytics literature by considering the serial mediation effect of data-driven culture and data quality. Few studies have tested sequential mediation effects relative to analytics capabilities on performance, therefore this approach is a strength of the paper [55, 56]. Maximizing data quality will necessitate maximization of various resources (e.g., analytics capabilities) to produce outcomes that otherwise do not exist. Third, this study extends the literature by considering the effect of an institutional construct in data-driven culture insofar that this would be related to the norms that bound the organization to IT change [57]. Fourth, data suggests that despite different funding mechanisms, data quality and data-driven culture are equally important for both public and private institutions.

This study looks at an important contemporaneous problem of how resource utilization can influence the organization from a fiscal perspective. This important problem facing higher education as well as many other business sectors.

5. Implications and Future Research

Although the study is theoretically grounded, there are some limitations. The sample population exists of

a subset of institutions of higher education and is limited to those members of Educause who participated in data collection surveys. Another limitation is the varying degree of the level of IT adoption and the degree of customization by institution. Customization may be both an obstacle and an opportunity. With any study, there are limitations, but there are also important implications, which emerge from this research.

Insight on mechanisms to improve data quality and efficiencies may lead to a more sustainable future in higher education. Maximization of these resources will help institutions build sustainable competitive advantage and RBT will help explain differences between institutions. Organizational performance as measured by net revenue per student would be a “standardized” proxy for efficiency as this metric is directly related to the use of core revenues and core expenses as defined by the education sector. These results also suggest future research into differences between private and public schools’ perceptions of impact between the data-driven culture and data quality relationship. This could provide meaningful insight to a fundamental difference between public and private institutions aside from the funding mechanisms. Considering that size exhibited a diminishing effect on organizational performance, it would seem that future research into this phenomena of whether larger institutions may have greater difficulty aligning data quality and data culture would have also have implications for practitioners.

6. References

- [1] NCES. *National Center for Education Statistics*. 2017 [cited 2017 July 1]; Available from: <https://nces.ed.gov/ipeds/trendgenerator/>.
- [2] Rabovsky, T.M., *Accountability in higher education: Exploring impacts on state budgets and institutional spending patterns*. Journal of Public Administration Research and Theory, 2012. **22**(4): p. 675-700.
- [3] Valbrun, M. *Discount Rates Hit Record Highs*. 2019 [cited 2020 January 7]; Available from: insidehighered.com/news/2019/05/10/nacubo-report-shows-tuition-discounting-trend-continuing-unabated.
- [4] Hildreth, B. *U.S. Colleges Are Facing a Demographic and Existential Crisis*. Huffington Post, 2017.
- [5] Grajek, S., *Divest, Reinvest, and Differentiate*, E.I.I. Panel, Editor. 2016, http://er.educause.edu/~media/files/articles/2016/1/er_m1611.pdf.
- [6] Tarafdar, M. and S.R. Gordon, *Understanding the influence of information systems competencies on process innovation: A resource-based view*. Journal of Strategic Information Systems, 2007. **16**(4): p. 353-392.
- [7] Mata, F.J., W.L. Fuerst, and J.B. Barney, *Information Technology and Sustained Competitive Advantage: A Resource-Based Analysis*. MIS Quarterly, 1995: p. 487-505.
- [8] Wade, M. and J. Hulland, *Review: the resource-based view and information systems research: review, extension, and suggestions for future research*. MIS Quarterly, 2004. **28**(1): p. 107-142.
- [9] Ard, T., J. Hurley, and S. Menditto, *Strategic Financial Planning & Analytics*. 2016, National Association of College and University Business Officers.
- [10] Manyika, J., et al., *Big data: The next frontier for innovation, competition, and productivity*. 2011.
- [11] Gupta, M. and J.F. George, *Toward the development of a big data analytics capability*. Information & Management, 2016. **53**(8): p. 1049-1064.
- [12] Reinitz, B.T., *Why Data and Analytics Matter - and why the Matter Now*. 2019, Educause: Report from the EDUCAUSE/NACUBO/AIR 2019 Enterprise Summit.
- [13] Barney, J. and W. Hesterly, *VRIO framework, strategic management and competitive advantage*. 2010, New Jersey: Pearson.
- [14] Sambamurthy, V., A. Bharadwaj, and V. Grover, *Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms*. MIS Quarterly, 2003. **27**(2): p. 237-263.
- [15] Campbell, C. and T. Fogarty, *Behind the Curve: Higher Education's Efforts to Implement Advanced Information Systems*. Journal of Emerging Technologies in Accounting, 2018. **15**(2): p. 77-91.
- [16] Norris, D. and L. Baer, *Building organizational capacity for analytics*. Educause Learning Initiative, 2013: p. 7-56.
- [17] Lyytinen, K. and V. Grover, *Management Misinformation Systems: A Time to Revisit?* Journal of the Association for Information Systems, 2017. **18**(3): p. 206.
- [18] Cao, G., Y. Duan, and G. Li, *Linking Business Analytics to Decision Making Effectiveness: A Path Model Analysis*. IEEE Transactions on Engineering Management, 2015. **62**(3): p. 384-395.
- [19] Grover, V., et al., *Creating Strategic Business Value from Big Data Analytics: A Research Framework*. Journal of Management Information Systems, 2018. **35**(2): p. 388-423.
- [20] Peteraf, M.A. and J.B. Barney, *Unraveling the resource-based tangle*. Managerial and Decision Economics, 2003. **24**(4): p. 309-323.
- [21] Makadok, R., *Toward a synthesis of the resource-based and dynamic-capability views of rent creation*. Strategic Management Journal, 2001. **22**(5): p. 387-401.
- [22] Ravichandran, T., C. Lertwongsatien, and C. Lertwongsatien, *Effect of information systems resources and capabilities on firm performance: A resource-based perspective*. Journal of management information systems, 2005. **21**(4): p. 237-276.
- [23] Bhatt, G.D. and V. Grover, *Types of Information Technology Capabilities and Their Role in Competitive Advantage: An Empirical Study*. Journal of Management Information Systems, 2005. **22**(2): p. 253-277.
- [24] Alavi, M. and D.E. Leidner, *Knowledge management and knowledge management systems: Conceptual*

- foundations and research issues*. MIS Quarterly, 2001: p. 107-136.
- [25] Kwon, O., N. Lee, and B. Shin, *Data quality management, data usage experience and acquisition intention of big data analytics*. International Journal of Information Management, 2014. **34**(3): p. 387-394.
- [26] Nikpour, A., *The impact of organizational culture on organizational performance: The mediating role of employee's organizational commitment*. International Journal of Organizational Leadership, 2017. **6**: p. 65-72.
- [27] Kim, G., B. Shin, and O. Kwon, *Investigating the Value of Sociomaterialism in Conceptualizing IT Capability of a Firm*. Journal of Management Information Systems, 2012. **29**(3): p. 327-362.
- [28] Dehning, B. and V.J. Richardson, *Returns on Investments in Information Technology: A Research Synthesis*. Journal of Information Systems, 2002. **16**(1): p. 7-30.
- [29] Tarhini, A., H. Ammar, and T. Tarhini, *Analysis of the critical success factors for enterprise resource planning implementation from stakeholders' perspective: A systematic review*. International Business Research, 2015. **8**(4): p. 25.
- [30] Chen, et al., *IT capability and organizational performance: the roles of business process agility and environmental factors*. European Journal of Information Systems, 2014. **23**(3): p. 326-342.
- [31] Hongjiang, X., et al., *Data quality issues in implementing an ERP*. Industrial Management & Data Systems, 2002. **102**(1): p. 47-58.
- [32] Kiron, D., et al., *Analytics: The widening divide*. MIT Sloan Management Review, 2012. **53**(2): p. 1.
- [33] Martinsons, M.G. and P.K. Chong, *The Influence of Human Factors and Specialist Involvement on Information Systems Success*. Human Relations, 1999. **52**(1): p. 123-152.
- [34] Ke, W. and K.K. Wei, *Organizational culture and leadership in ERP implementation*. Decision Support Systems, 2008. **45**(2): p. 208-218.
- [35] Elbashir, M.Z., et al., *Enhancing the Business Value of Business Intelligence: The Role of Shared Knowledge and Assimilation*. Journal of Information Systems, 2013. **27**(2): p. 87-105.
- [36] Hayes, A.F., *Introduction to Mediation, Moderation, and Conditional Process Analysis*. 2013, New York, NY: The Guilford Press. 418.
- [37] LaValle, S., et al., *Analytics: The new path to value*. MIT Sloan Management Review, 2010. **52**(1): p. 1-25.
- [38] Chen and Oztekin, *A hybrid data envelopment analysis approach to analyse college graduation rate at higher education institutions*. INFOR, 2017. **55**(3): p. 188-210.
- [39] Pike, G.R. and S.S. Graunke, *Examining the Effects of Institutional and Cohort Characteristics on Retention Rate*. Research in Higher Education, 2015. **56**(2): p. 146-165.
- [40] Hurlburt, S. and R.J. Kirshstein, *Spending: Where Does the Money Go? A Delta Data Update, 2000-2010*. Delta Cost Project at American Institutes for Research, 2012.
- [41] Wei, Z., O. Gutierrez, and K. Mathieson, *Information Systems Research in the Nonprofit Context: Challenges and Opportunities*. Communications of the Association for Information Systems, 2010. **27**: p. 1-12.
- [42] Tandberg, D.A. and N.W. Hillman, *State Higher Education Performance Funding: Data, Outcomes, and Policy Implications**. Journal of Education Finance, 2014. **39**(3): p. 222-243.
- [43] Muthén, B. and D. Kaplan, *A comparison of some methodologies for the factor analysis of non-normal Likert variables*. British Journal of Mathematical and Statistical Psychology, 1985. **38**(2): p. 171-189.
- [44] MacCallum, R.C., et al., *Sample size in factor analysis*. Psychological Methods, 1999. **4**(1): p. 84-99.
- [45] MacKenzie, S.B., P.M. Podsakoff, and C.B. Jarvis, *The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions*. Journal of Applied Psychology, 2005. **90**(4): p. 710.
- [46] Hair Jr, J.F., et al., *Multivariate Data Analysis*. 7th ed. 2010, Upper Saddle River, NJ: Prentice Hall.
- [47] Brown, T.A., *Confirmatory Factor Analysis for Applied Research*. 2015, New York, NY: The Guilford Press. 429.
- [48] Hu, L. and P.M. Bentler, *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives*. Structural Equation Modeling: A Multidisciplinary Journal, 1999. **6**(1): p. 1-55.
- [49] Hooper, D., J. Coughlan, and M. Mullen, *Structural equation modelling: Guidelines for determining model fit*. Articles, 2008: p. 2.
- [50] Podsakoff, P.M., S.B. MacKenzie, and N.P. Podsakoff, *Sources of method bias in social science research and recommendations on how to control it*. Annual Review of Psychology, 2012. **63**: p. 539-569.
- [51] Bagozzi, R.P., *Measurement and Meaning in Information Systems and Organizational Research: Methodological and Philosophical Foundations*. MIS Quarterly, 2011. **35**(2): p. 261-292.
- [52] O'Brien, R.M., *A caution regarding rules of thumb for variance inflation factors*. Quality & Quantity, 2007. **41**(5): p. 673-690.
- [53] Gold, A.H., A. Malhotra, and A.H. Segars, *Knowledge management: An organizational capabilities perspective*. Journal of management information systems, 2001. **18**(1): p. 185-214.
- [54] Drucker, P.F., *Work and tools*. Technology and Culture, 1959. **1**(1): p. 28-37.
- [55] Huertas-Valdivia, I., F.J. Llorens-Montes, and A. Ruiz-Moreno, *Achieving engagement among hospitality employees: A serial mediation model*. International Journal of Contemporary Hospitality Management, 2018. **30**(1): p. 217-241.
- [56] Jones, R.A., N.L. Jimmieson, and A. Griffiths, *The impact of organizational culture and reshaping capabilities on change implementation success: The mediating role of readiness for change*. Journal of Management Studies, 2005. **42**(2): p. 361-386.
- [57] Weerakkody, V., Y.K. Dwivedi, and Z. Irani, *The diffusion and use of institutional theory: a cross-disciplinary longitudinal literature survey*. Journal of Information Technology, 2009. **24**(4): p. 354-368.