Big Data Analytics Capability and Governmental Performance:
An Empirical Examination

Guido Ongena, HU University of Applied Sciences, Utrecht, The Netherlands*
Arjen Davids, Municipality of Berg en Dal, The Netherlands

ABSTRACT

Although governments are investing heavily in big data analytics, reports show mixed results in terms of performance. Whilst big data analytics capability provided a valuable lens in business and seems useful for the public sector, there is little knowledge of its relationship with governmental performance. This study aims to explain how big data analytics capability led to governmental performance. Using a survey research methodology, an integrated conceptual model is proposed highlighting a comprehensive set of big data analytics resources influencing governmental performance. The conceptual model was developed based on prior literature. Using a PLS-SEM approach, the results strongly support the posited hypotheses. Big data analytics capability has a strong impact on governmental efficiency, effectiveness, and fairness. The findings of this paper confirmed the imperative role of big data analytics capability in governmental performance in the public sector, which earlier studies found in the private sector. This study also validated measures of governmental performance.

KEYWORDS

Big Data Analytics Capability, Effectiveness, Efficiency, Fairness, Governmental Performance, Structural Equation Modeling

INTRODUCTION

Propelled by increased accessible infrastructure and computing power, and the acquisition of more volumes of data accumulate into big data it is thought to be one of the most valuable strategic business sources in the coming years (McAfee & Brynjolfsson, 2012). This impact of big data analytics is potentially noticeable in a wide variety of sectors. Many scholars stipulate the future importance and value creation of big data analytics in hospitality (Horng, Lio, Chou, Yu, & Hu, 2022), healthcare (Yu, Zhao, Liu, & Song, 2021), retail (Santoro, Fiano, Bertoldi, & Ciampi, 2019), circular economy (Kristoffersen, Mikalef, Blomsma, & Li, 2021), food industry (Chakraborty, Rana, Khorana, Singu, & Luthra, 2022), and supply chain (Gopal, Rana, Krishna, & Ramkumar, 2022). The same rule applies

DOI: 10.4018/IJEGR.321638 *Corresponding Author

to the public sector. Big data analytics have potentially many advantages, in terms of smart services, intelligent adaptive forms and predictive service delivery to its citizen, if they are used efficiently and effectively (Merhi & Bregu, 2020), also for smaller governments like municipalities (Milakovich, 2012). For this reason, governments are investing heavily in (big) data analytics (Gartner, 2019).

To reap the benefits of big data analytics it is imperative to gain an understanding of how organizations build big data analytics capabilities. This is important since we know from previous research, that adopted the theoretical lens of the resource-based theory (RBT), that organizations achieve competitive advantage by building capabilities, which in turn are created by combining and deploying several resources. Based on the RBT, Gupta and George (2016) suggests that organizations should focus on creating a big data capability to achieve sustainable competitive advantage by integrating its tangible resources (e.g., data), human resources (e.g., technical skills) and intangibles resources (e.g., data-driven culture). They juxtaposed these three resources that together build a big data analytics capability. Studies found empirical evidence that these resources contribute to the organization's performance (Wamba, et al., 2017; Ferraris, Devalle, & Couturier, 2019; Mikalef, Krogstie, Pappas, & Pavlou, 2020). A big data analytics capability has thus been shown to positively impact business performance in studies on business organizations.

Unfortunately, how big data analytics capability creates value for the public sector is not sufficiently empirically assessed in the extant literature. Most reports on the value of big data to date have been from consultancy firms (e.g., EY, 2021), and conceptual studies (e.g., Merhi & Bregu, 2020) that lack empirical theoretical insight. As a result, there is limited understanding of how organizations should approach their big data initiatives and scarce empirical support to back-up the claim that these investments result in any measurable administrative value. This study extends the stream of research on big data analytics capability and organizational performance by examining factors that contribute to improved governmental performance because of investments in big data analytics. More specifically, the study aims to examine the following research question:

Does a big data analytics capability result in governmental performance gains?

This study addresses this research question by developing a conceptual model to study the big data analytics capability in relation to governmental performance (i.e., fairness, efficiency, and effectiveness). In doing so, concepts from the big data analytics capabilities literature are adopted (Gupta & George, 2016) that stem from business management research and combine these with governmental performance based on measures developed in public administration research (Kim, 2005; Brewer & Selden, 2000). To operationalize this, a survey was developed based on previous measures and was distributed to municipalities in the Netherlands. A robust quantitative analysis was performed by adopting a structural equation modelling approach. Our results indicate an imperative role of big data analytics capabilities as it significantly affects governmental performance.

The rest of the paper is structured as follows. In the literature review section that follows, the relevant academic literature is described in this study highlighting the need to look at the big data analytics capabilities of public organizations. In Section 3 the research model is introduced as are the corresponding research hypotheses. In Section 4, the followed research method is presented to actualize the study's objectives, followed in Section 5 by the empirical analysis and the outcomes that include an assessment of the measurement model and the structural model. Section 6 concludes by discussing the findings from a research and practical standpoint and outlining some key limitations that underpinned this study.

LITERATURE REVIEW

Big Data Analytics Capability

Big data analytics is generally characterized by its Three V's that have emerged as a common framework to describe big data (Chen, Chiang, & Storey, 2012): volume (refers to the large magnitude of data), variety (refers to its structural heterogeneity), velocity (refers to its high speed of generation). Although relevant, the v's by which big data is characterized are rather limited in scope as they mainly pertain to technical aspects. Gupta and George (2016) embarked on this shortcoming and developed a more organization-wide capability framework for big data analytics based on the resource-based theory. This theory stems from strategic management literature, which has proven its value over the years and is one of the most prominent and powerful theories for understanding organizations (Barney, Ketchen, & Wright, 2011).

Drawing on this theory a big data analytics capability was proposed consisting of tangible, human, and intangible resources (Gupta & George, 2016). First, tangible resources pertain much to the aforementioned characteristics of big data. Access and integration of internal and external data (Zhao, Fan, & Hu, 2014) is an imperative tangible resource. In addition to the data itself, organizations must possess technological and physical infrastructure requirements that allow for the efficient use of data. Besides data and technology, organizations need to make adequate investments in their big data initiatives in terms of time and money. This is labelled as the basic resources (Gupta & George, 2016) and is also the most fundamental variable for innovation using open data (Dwivedi, et al., 2017). Second, human resources are formulated in terms of skills. Two types of skills are defined. On the one hand, skills pertain to people who have technical competences to work with data. Skills of this type include exploratory data analysis, statistics, machine learning and programming. On the other hand, skills are related to people who possess management competences. It is imperative that managers advocate the use of big data throughout the organization and seek value creation through this usage. Managers ensure mutual trust and a good working relationship between big data managers and other functional managers (Gupta & George, 2016). Third and last, intangible resources are reflected by a data-driven culture that enables data-driven decision-making, rather than following intuition, by managers at any level in the organizations, and organizational learning. Organizational learning suggests organizations that have developed capabilities to explore, accumulate, share and transform knowledge possess a key inventory of valuable knowledge, very useful when validating and contextualizing the results obtained from big data (Lozada, Arias-Perez, & Perdomo-Charry, 2019).

Governmental Performance

A myriad of researchers examined the black box of governmental performance over the years. Rainey and Steinbauer (1999), for instance, developed a theory that posited that organizational performance was affected by political authorities, agency autonomy in refining and implementing its mission, high mission valence, a strong, mission-oriented culture, and certain leadership behaviours. These determinants were empirically tested which confirmed most of the hypotheses (Brewer & Selden, 2000). Later studies found additional determinants that positively affect organizational performance, especially from a human-resource perspective. For instance, Giauque, Anderfuhren-Biget, and Varone (2013) found several intrinsic motivators (e.g., fairness, job enrichment, individual appraisal, and professional development) that contribute to organizational performance. Similarly, research showed a positive impact of family-friendly work practices (Ko, Hur, & Smith-Walker, 2013).

Although Rainey and Steinbauer (1999) already provided a valuable stepping-stone with their proposed theory of effective government organization two decades ago, much attention was paid to the determinants of governmental performance, but less to the concept itself. The measurement of performance by governmental bodies is thus still a youthful and under-investigated field of research (Andrews, Boyne, Moon, & Walkter, 2010). Therefore, there is hitherto no consensus on how to measure governmental performance. The proposed measures of organizational performance based on the perceptions of the organization's members by Brewer and Selden (2000) are to date the most comprehensive and theoretically founded measurement instrument, that is also successfully used by other scholars (e.g., Kim, 2005). Their measurement comprises two organizational foci dimensions (i.e., internal and external) and three administrative value dimensions (i.e., efficiency, effectiveness, and fairness). Table 1 shows the different dimensions.

Table 1.
Dimensions of governmental performance (adopted from Brewer and Selden, 2000)

		Administrative Values		
		Efficiency	**Effectiveness**	**Fairness**
Organizational Focus	Internal	Internal Efficiency (e.g., low performance costs)	Internal Effectiveness (e.g., high productivity)	Internal Fairness (e.g., equitable treatment for employees)
	External	External Efficiency (e.g., prompt business relations)	External Effectiveness (e.g., goal-oriented)	External Fairness (e.g., equitable services)

Big Data Analytics and Governmental Performance

Early research centred on IT capabilities and performance showed a positive relationship. Firms with high IT capability tend to outperform firms on a variety of performance measures (Bharadwaj, 2000). More specific to big data analytics capabilities, studies show a broad consensus that big data analytics enables firms, thus in the context of commercial organizations, to attain a state of competitive advantage by strengthening intermediate organizational capabilities (Wamba, et al., 2017; Ferraris, Devalle, & Couturier, 2019; Mikalef, Krogstie, Pappas, & Pavlou, 2020; Chatterjee, Rana, & Dwivedi, 2021). The value of investing in big data analytics capability in the private sector is thus clearly reflected in the literature. Although empirical studies focused on the performance gains of developing a big data analytics capability in the governmental context are rather scarce, some research has demonstrated a positive overall association (Milakovich, 2012). Overall, big data analytics capability thus can be considered both a resource and a capability that can enable efficient and effective business operations.

RESEARCH MODEL AND HYPOTHESES

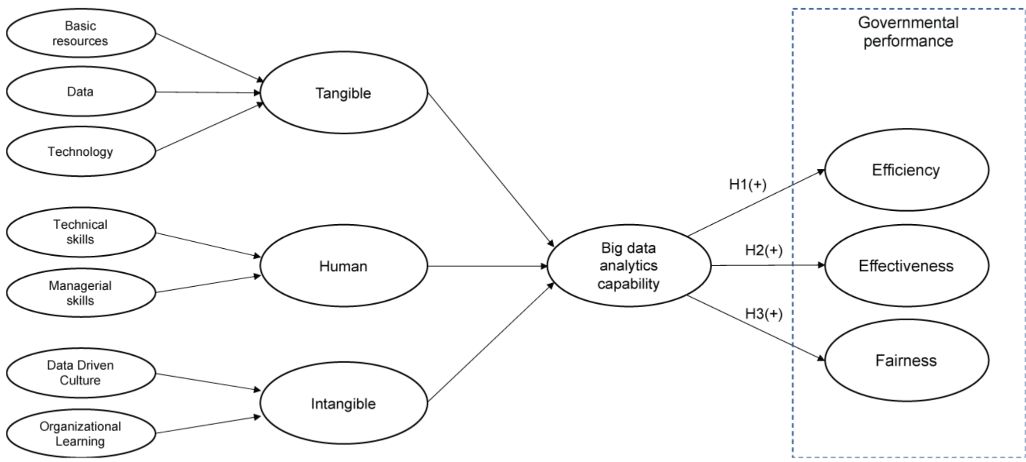
Drawing on the literature, this study proposes the research model shown in Fig. 1. Similar to previous literature on big data analytics capabilities (Gupta & George, 2016; Mikalef, Boura, Lekakos, & Krogstie, 2019; Mikalef, Krogstie, Pappas, & Pavlou, 2020; Lozada, Arias-Perez, & Perdomo-Charry, 2019), this study proposes big data analytics capabilities as a third-order, hierarchical model manifested in three second-order constructs: tangible resources, human resources, intangible resources. Figure 1 illustrates the conceptual model. The main proposition is that governmental performance (i.e., effectiveness, efficiency, fairness) varies depending on the level of the big data analytics capability.

Foremost, big data analytics initiatives in governments seek to effect efficiencies (Morabito, 2015). Big data analytics can boost efficiency by reducing the number of inputs necessary for providing the current service level and/or producing the actual output level (Rogge, Agasisti, & De Witte, 2017). This promise to enhance operational efficiency triggered governments of leading ICT countries to initiate big data application projects (Kim, Trimi, & Chung, 2014). Practice-based reports also stipulate the potential savings by governments in utilizing big data analytics. By making smarter decisions about how departments are organized and what work gets prioritized and making more efficient use of resources (Ubaldi, Van Ooijen, & Welby, 2019), the direct cost of government operations can be reduced (Manyika, et al., 2011) that can also be experienced by citizens in their day to day lives (Welby, 2019). Based on this alleged potential of big data analytics to enhance efficiency, the following hypothesis is posited:

H1: Big data analytics capabilities have a positive effect on the efficiency of the organization.

As Rogge, Agasisti and De Witte (2017, p. 275) stipulate, it is important to make a distinction between efficiency and effectiveness. In contrast to efficiency, evaluations of effectiveness focus on

Figure 1.
Research model



the link between the outputs and the outcomes. In relation to big data analytics, this pertains to the quality of decisions. Literature in the private sector suggests that big data analytics allegedly improves the quality of these decisions (McAfee & Brynjolfsson, 2012) which is empirically supported by Ghasemaghaei, Ebrahimi and Hassanein their study (2018). This potential of improved decision quality by big data analytics is also addressed for policy, as it can improve policy analysis and decision-making by policymakers (Pencheva, Esteve, & Mikhaylov, 2020; Ubaldi, Van Ooijen, & Welby, 2019). Based on this argumentation the following hypothesis is thus formulated:

H2: Big data analytics capabilities have a positive effect on the effectiveness of the organization.

Where efficiency and effectiveness are similar goals in both the private as well as the public sector, fairness is a goal that explicitly is an administrative value. This includes preventing unfair discrimination against individuals and preventing the unfair exploitation of individuals (Bannister & Connolly, 2014). Much literature stipulates that big data analytics poses considerable risks to these values (e.g., Favaretto, De Clercq and Elger, 2019). Although the literature suggests that adequate big data analytics governance could overcome any unexpected incident that might occur due to the deployment of an inappropriate usage (Rana, Chatterjee, Dwivedi, & Akter, 2022), the algorithms are designed by humans and increasingly learn by observing human behaviour through data. Therefore, they tend to adopt the biases that exist in the analysed data (Murreddu, Schmeling, & Kanellou, 2020). Conversely, big data analytics could also be an important instrument to prevent inequality and discrimination. Big data analytics is said to promote objectivity in classification and profiling because decisions are made by a formal, objective, and constant algorithmic process with a more reliable empirical foundation than the human decision-making (Barocas & Selbst, 2016), omitting human biases. An eminent and often-used example is that big data analytics has the potential to spot fraud or corruption (Cunningham, McMillan, O'Rourke, & Schweikert, 2018; Ubaldi, Van Ooijen, & Welby, 2019; Munné, 2016). On this note, this study postulates the following hypothesis:

H3: Big data analytics capabilities have a positive effect on the fairness of the organization.

RESEARCH METHOD

Development of Measurement Instrument

In developing the measurement instrument measures from existing literature were adopted, which represents a legitimate way to create an initial scale, simultaneously ensuring content validity. Table 6 in Appendix A lists the operationalization of all the constructs. The measures concerning big data analytics capabilities consisted of 16 items. In line with the study by Gupta and George (2016), the tangible resources were modelled as a type IV second-order construct, meaning that both first-order and second-order constructs are formative. Human and intangible resources were modelled as a type II second-order construct, which means that the first-order constructs are reflective and the second-order are formative. The items were measured on a 7-point Likert scale (from strongly disagree to strongly agree).

The 12 items as suggested by Kim (2005) were used to measure governmental performance. Each part of the governmental performance (i.e., effectiveness, efficiency, fairness) entailed 4 items. The internal and external perspective was equally divided into 2 items each. The items were measured on a 7-point Likert scale (from strongly disagree to strongly agree).

Sampling and Data Collection

The data used in this research was collected through a self-administered online questionnaire targeted at employees of municipalities during the autumn of 2021. The data was collected in the Netherlands, which consists of 352 municipalities. Municipalities were contacted directly via email. A total of 318 respondents started to complete the survey, with 120 providing complete responses. Table 2 shows the distribution of size-classes in terms of number of employees and years of experience with data. As can be seen in the same table, the experience of the respondents is somewhat skewed to both ends of the spectrum. Most respondents are experienced with data (more than 4 years). This is not surprising as previous research found that the organizational capabilities of Dutch government organizations are quite well developed, on average, which could lead organizations to believe they are ready to start using big data (Klievink, Cunningham, & de Bruijn, 2017).

Non-response bias was assessed. Non-response bias refers to a situation in which people who do not respond to a questionnaire may bias the research results. To determine whether there was any non-response bias in this study's sample, the profile of the respondents was compared with those on the mailing list that was collected for each municipality, such as size. In addition, the nonresponse approach follows the suggested wave analysis by Armstrong and Overton (1977), who suggested that late respondents are more likely to resemble non-respondents than to resemble early respondents. The first and last waves of respondents on all the variables are compared, which treats late respondents as a proxy of non-respondents. No statistically significant differences were found ($p < 0.01$). Hence, this shows that there is no critical degree of non-response bias with the used data.

To assess common method bias, this study first used procedural controls (ex ante*)* during the design of the survey (Podsakoff, MacKenzie, & Podsakoff, 2012). Respondents were assured that all the information they provided would remain completely anonymous and confidential and stipulated and that any analysis would be done on an aggregate level solely for research purposes. Also, clear instructions were provided to avoid complex and ambiguous items. The latter was done through pre-testing by survey experts, which subsequently refined the formulation of the questions were further and eliminated any repeated or similarly sounding items. In addition, the Harman's one factor test was employed as a statistical control (ex post). The results show that 35.7% of the variance was explained by one single factor. Since the explained variance should not exceed 50% or more before substantial concern about inflated relationships would arise. To summarize, these steps assured that common method bias was not an issue with the collected data.

Table 2.
Sample characteristics

		Frequency (N=120)	Percentage (%)
Gender	Male	82	68.3%
	Female	31	25.8%
	Don't want to say	7	5.8%
Number of employees	1-50	7	5.8%
	51-100	5	4.2%
	101-1,000	70	58.3%
	More than 1,000	38	31.7%
Big data experience	Less than 1 year	38	31.7%
	1 - 2 years	16	13.3%
	2 - 3 years	10	8.3%
	3 - 4 years	11	9.2%
	More than 4 years	45	37.5%

RESULTS

SmartPLS (Ringle, Wende, & Becker, 2015) is used to analyse the data variance-based structural equation modelling (PLS-SEM). Reasons for the use of PLS-SEM include 1) the research model includes one or more formatively measured constructs, 2) the emphasis of this study is on prediction and theory development rather than theory testing, 3) less demand on the distribution of the variables (Hair, Risher, Sarstedt, & Ringle, 2019).

Assessment of the Measurement Model

As suggested by literature (Hair, Risher, Sarstedt, & Ringle, 2019), different assessment criteria are used for the evaluation of reflective constructs and formative constructs.

For reflective latent constructs, the reliability, convergent validity and discriminant validity are assessed. Internal consistency reliability of the measures is shown by the values of the Cronbach's alphas and composite reliabilities as they are above the threshold of 0.70 (see Table 3), suggesting a satisfactory level of construct reliability. Convergent validity is also established since the average variance extracted is at least 0.50 (see Table 3) for all measures which demonstrate sufficient results. The discriminant validity is assessed in three ways. First, the Fornell–Larcker criterion is used which states that correlations of the construct with other constructs should be lower than the square root of average variance extracted of the construct (Fornell & Larcker, 1981). The values of the latter are shown in the off-diagonal of Table 3 and the correlations below. All values comply with this criterion which suggests discriminant validity. Second, the loadings of the reflective indicators on their latent constructs should be notably larger than their cross-loadings at the item level. As can be seen in Table 7 in Appendix B this holds true for all measurement items except for two items from *efficiency*, namely EC1 and EC2. Similarly, the construct does not fully comply with heterotrait-monotrait (HTMT) ratio, the third way to assess discriminant validity. To clearly discriminate between two factors, the HTMT should be smaller than 0.90 (Henseler, Hubona, & Ray, 2016). All the obtained correlations (see Table 8 in Appendix C) comply with that criterion demonstrating discriminant validity except for *efficiency*. The measures used seem not discriminating against the construct of *effectiveness*. However, the items were retained for two reasons: 1) one of the three, the Fornell–Larcker criterion, showed discriminant validity, and 2) more important, the full set of measures is adopted from prior literature based on theoretical groundings. It is therefore deemed not necessary to remove any items. To summarize, the results suggest that first-order reflective measures are valid to work with and support the appropriateness of all items as suited indicators for their respective constructs.

The research model entails three first-order formative latent constructs (i.e., basic resources, data, and technology). Also, the high-order constructs are modelled as formative constructs. First, to examine potential multicollinearity issues, the variance inflation factor (VIF) values are evaluated. The VIF values should be below 3.3 (Diamantopoulos & and Siguaw, 2006). Table 4 presents the VIF values of the measures used, which show satisfactory values below this threshold. Hence, this suggests that collinearity was not a major issue in the study. Second, the measures' weights and respective significance levels are assessed. All the weights present satisfactory significance levels. The outer loadings also exceed the threshold of 0.5 (Hair, Risher, Sarstedt, & Ringle, 2019).

Assessment of the Structural Model

Using the PLS algorithm, the explanatory and predictive power were obtained. Also, the size of path coefficients, and via the bootstrap approach the significance, was obtained as shown in Figure 2.

The explanatory power (R^2) is examined. R^2 measures the variance which is explained in each of the endogenous constructs. The structural model explains 31.8% of the variance in the first endogenous construct in the research model, efficiency ($R^2 = 0.318$). In addition, the model explains 33.0% of the variance for effectiveness ($R^2 = 0.330$). These coefficients of determination represent a moderate explanatory power (Hair, Risher, Sarstedt, & Ringle, 2019). The last, endogenous constructs, presents weak explanatory power by big data analytics capability as it explains only 5% of the variance. ($R^2 =0.050$).

In addition to the in-sample explanatory power measured by R^2, the out-of-sample predictive power of the model is assessed by conducting the PLSpredict procedure (Shmueli, Ray, Velasquez Estrada, & Shatla, 2016). Table 5 shows that, for the majority, the prediction error values of the PLS-SEM, root mean squared error (RMSE) or mean absolute error (MAE) are lower in comparison with the values of a linear regression model (LM). This indicates that the model has a medium to high out-of-sample power (Shmueli, Ray, Velasquez Estrada, & Shatla, 2016).

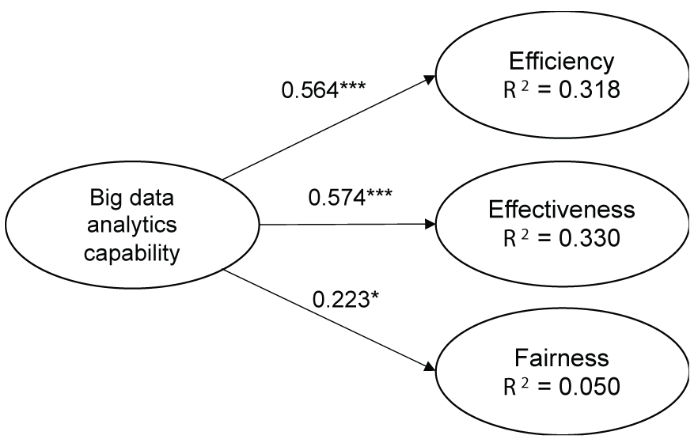
Table 3.
Measures of reliability, convergent, and discriminant validity of reflective constructs

		1	2	3	4	5	6	7	8	9	10
1	Basic resources	n/a									
2	Data	0.576	n/a								
3	Technology	0.573	0.603	n/a							
4	Technical skills	0.497	0.658	0.638	*0.959*						
5	Management skills	0.495	0.545	0.501	0.488	*0.879*					
6	Data driven culture	0.594	0.470	0.478	0.527	0.556	*0.861*				
7	Organizational learning	0.500	0.522	0.376	0.613	0.586	0.597	*0.932*			
8	Effectiveness	0.355	0.371	0.339	0.479	0.514	0.532	0.580	*0.780*		
9	Efficiency	0.358	0.360	0.382	0.561	0.499	0.452	0.620	0.724	*0.728*	
10	Fairness	0.019	0.141	0.030	0.175	0.368	0.188	0.427	0.604	0.483	*0.801*
Cronbach's Alpha		n/a	n/a	n/a	0.912	0.705	0.825	0.849	0.790	0.710	0.827
Composite Reliability		n/a	n/a	n/a	0.958	0.871	0.896	0.930	0.860	0.813	0.877
Average Variance Extracted		n/a	n/a	n/a	0.919	0.772	0.742	0.869	0.608	0.530	0.641

Table 4.
Measures of validation for formative constructs

Construct	Measures	VIF	Weight	Significance	Loading
Basic resources	BR1	2.561	0.549	p<0.001	0.948
	BR2	2.561	0.511	p<0.001	0.939
Data	D1	1.633	0.243	p<0.05	0.760
	D2	1.633	0.831	p<0.001	0.982
Technology	T1	1.746	0.417	p<0.001	0.772
	T2	1.496	0.577	p<0.001	0.856
	T3	2.269	0.222	p<0.05	0.827
Tangible	Basic resources	1.645	0.482	p<0.001	0.873
	Data	1.617	0.364	p<0.001	0.814
	Technology	1.699	0.345	p<0.001	0.821
Human skills	Technical skills	1.312	0.656	p<0.001	0.900
	Management skills	1.312	0.500	p<0.001	0.820
Intangible	Data driven culture	1.554	0.644	p<0.001	0.926
	Organizational learning	1.554	0.471	p<0.001	0.856
Big data analytics capability	Tangible	2.494	-0.369	p<0.001	0.585
	Human skills	3.125	0.718	p<0.001	0.896
	Intangible	2.247	0.626	p<0.001	0.914

Figure 2.
Estimated relationships of the structural model



The path coefficients, and significance of estimates (t-statistics), are obtained by performing a bootstrap analysis with 5000 resamples. The results reveal a significant influence of big data analytics capability on efficiency (β=0.564, t=8.403, p<0.001), which means that H1 is supported. Similarly, big data analytics capability significantly affects the effectiveness of municipalities (β=0.574, t=9.481,

Table 5.
PLSpredict assessment of manifest variables

	PLS-SEM		LM		PLS<LM	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
EV1	1.386	1.035	1.528	1.151	yes	yes
EV2	1.251	0.952	1.214	0.931	no	no
EV3	1.303	1.091	1.370	1.125	yes	yes
EV4	1.101	0.821	1.101	0.849	no	yes
EC1	1.324	1.024	1.463	1.109	yes	yes
EC2	1.273	1.016	1.408	1.135	yes	yes
EC3	1.303	1.047	1.366	1.081	yes	yes
EC4	1.303	1.029	1.409	1.112	yes	yes
F1	1.102	0.834	1.143	0.872	yes	yes
F2	1.201	0.838	1.275	0.929	yes	yes
F3	1.057	0.811	1.061	0.857	yes	yes
F4	1.325	0.988	1.384	1.015	yes	yes

$p<0.001$). This result supports H2. The results also present a significant and positive relationship between big data analytics capability and fairness ($\beta=0.223$, $t=1.775$, $p<0.10$), thus supporting H3.

DISCUSSION

The present study has developed a big data analytics capability in the context of the public sector. By doing so the applicability of the RBT is extended from the private sector, which is to date the dominant examined sector (e.g., Mikalef, Krogstie, Pappas, & Pavlou, 2020), to the public sector. This is the first study to examine this big data analytics capability in relation to governmental performance. The results, as shown in Figure 2, support the claim of the positive impact of big data analytics capabilities on the efficiency of governments (H1). This result shows empirical support for hitherto anecdotal evidence regarding the impact of digital leadership on digital transformation, such as Rogge, Agasisti, and De Witte (2017). Another interesting finding, related to H2, is that big data analytics capability significantly affects the quality of decision making, thus the effectiveness of the organization. These results are consistent with previous research that empirically examined this relationship in the private sector (Ghasemaghaei, Ebrahimi, & Hassanein, 2018) and provides empirical evidence for literature that allegedly stipulated this relationship (Pencheva, Esteve, & Mikhaylov, 2020; Ubaldi, Van Ooijen, & Welby, 2019). Moreover, this study (H3) confirmed the suggested relationship between big data analytics capability and fairness. Although there is much scepticism on the use of big data analytics, by finding a positive significant result between the two constructs this study showed empirical evidence that developing a proper big data analytics capability enhances the fairness of the organization.

Theoretical Implications

The purpose of this research is to expand on the current body of knowledge regarding big analytics, with the aim of presenting a more thorough and knowledgeable account of how big data analytics capability generates value for the public sector. To achieve this goal, a resource-based view framework was employed in the construction of the proposed model. This framework is recognized as a useful managerial tool for identifying strategic resources that organizations can leverage to achieve sustainable business value. By bringing together various elements of big data analytics capability, which is already established in the literature, this study has developed a theoretical model aimed at enhancing governmental performance.

This research makes several contributions to the body of knowledge to explain governmental performance through big data analytics capabilities. In particular, this study contributes to theory in three ways. First, based on previous literature (e.g., Gupta and George, 2016) a conceptual model is developed sprouted from the resource-based theory on the relationship between big data analytics capability and governmental performance. As a myriad of previous literature addressed the contribution of big data analytics capability to create business value, there was a need to validate the conceptual model for the governmental domain. Second, this study contributed to the measurement of governmental performance. The measurement instrument developed by Kim (2005) was adopted. This study further validated the instrument by following a scientific approach with appropriate statistical indicators to confirm validity and reliability. This has led to the development of a scale useful for future governmental performance studies. It is asserted that this study has contributed to the existing literature on this topic Finally, it is worth noting that this work also extends the body of knowledge in the fields of research related to public values (Bannister & Connolly, 2014). This study addresses the suggested shortcoming of whether big data analytics increases fairness by yielding empirical support for the theoretical framework. This study makes an important contribution to this literature by presenting how big data analytics capability positively affects fairness. This is perhaps the first research that has highlighted the need of encompassing public values in the performance of public administration.

Managerial Implications

The outcomes of this study also present several interesting implications for practice. it is evident that investing in big data analytics capability can lead to significant improvements in the efficiency and effectiveness of municipalities. Thus, public sector organizations should prioritize the development and implementation of big data analytics capabilities to improve their operational efficiency and effectiveness. The findings strongly suggest that managers in governmental agencies must address big data analytics organization-wide. A narrow focus on the technical aspects of big data is too limited to adequately create value. One must pay attention to intangible aspects, such as a data-driven culture and organizational learning. Moreover, to enhance the value creation of big data analytics human skills play a central role. Training and educating personnel on a technical level is imperative. Additionally, management skills must be improved. It is also notable that, next to the impact on efficiency and effectiveness, big data analytics capability positively affects the fairness of the organization. Big data analytics thus can potentially support a fairer society, which is an imperative administrative value. The finding that big data analytics capability is positively related to fairness suggests that organizations can utilize this capability to enhance transparency and equity in decision-making processes. Therefore, public sector managers should consider the potential of big data analytics to promote fairness in their operations. In sum, the significance of these relationships underscores the need for public organizations to develop and implement robust data analytics capabilities to achieve better performance in the long run.

Limitations and Future Research Directions

While the findings of this study offer valuable insights into the impact of big data analytics capability on organizational performance in municipalities, there are several limitations to consider. Firstly, the study only examined the research model from the perspective of municipalities, which may limit the generalizability of the findings to national government agencies or public sector entities at other government levels. Future research should aim to investigate the applicability of the proposed model to these populations. Secondly, while reliability and convergent validity were established for the measurement of governmental performance, the study did not conclusively establish discriminant validity. Additionally, no measurement instrument currently exists that fully reflects administrative values such as inclusiveness and transparency. Therefore, future research should focus on developing a more rigorous and comprehensive instrument for measuring governmental performance. Lastly, the study used cross-sectional data, which may limit its ability to establish the stability of the findings over time. As such, it is recommended that future research employ longitudinal data to examine the stability of the results. This is particularly important given the skewness of the sample towards innovators. Overall, while this study provides valuable insights, these limitations should be kept in mind when interpreting the results and designing future research.

CONCLUSION

In light of the study's findings, it can be concluded that big data analytics capability plays a critical role in shaping governmental performance. Specifically, the results confirm a strong relationship between big data analytics capability and both effectiveness and efficiency. These findings highlight the potential of big data analytics capability to drive organizational performance and improve the delivery of public services in municipalities. Moreover, this study is the first to establish a positive relationship between big data analytics capability and fairness in governmental performance. This finding emphasizes the potential of big data analytics to promote transparency and equity in decision-making processes, thereby enhancing trust in public institutions and ultimately improving public satisfaction with the services provided. Overall, this study's results provide important insights into the role of big data analytics capability in shaping governmental performance, highlighting its potential to drive improvements in effectiveness, efficiency, and fairness. These findings have significant implications for public sector organizations seeking to improve their operations and better serve their constituents, underscoring the importance of investing in robust data analytics capabilities. However, further research is needed to address the limitations of the study and extend its findings to other government levels and public sector entities.

REFERENCES

Andrews, R., Boyne, G., Moon, M., & Walkter, R. (2010). Assessing Organizational Performance: Exploring Differences Between Internal and External Measures. *International Public Management Journal*, *13*(2), 105–129. doi:10.1080/10967491003766533

Armstrong, J., & Overton, T. (1977). Estimating nonresponse bias in mail surveys. *JMR, Journal of Marketing Research*, *14*(3), 396–402. doi:10.1177/002224377701400320

Bannister, F., & Connolly, R. (2014). ICT, public values and transformative government: A framework and programme for research. *Government Information Quarterly*, *31*(1), 119–128. doi:10.1016/j.giq.2013.06.002

Barney, J., Ketchen, D. Jr, & Wright, M. (2011). The Future of Resource-Based Theory: Revitalization or Decline? *Journal of Management*, *37*(5), 1299–1315. doi:10.1177/0149206310391805

Barocas, S., & Selbst, A. (2016). Big Data's disparate impact. *California Law Review*, *104*(3), 671–732.

Bharadwaj, A. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *Management Information Systems Quarterly*, *24*(1), 169–196. doi:10.2307/3250983

Brewer, G., & Selden, S. (2000). Why Elephants Gallop: Assessing and Predicting Organizational Performance in Federal Agencies. *Journal of Public Administration: Research and Theory*, *10*(4), 685–712. doi:10.1093/oxfordjournals.jpart.a024287

Chakraborty, D., Rana, N., Khorana, S., Singu, H., & Luthra, S. (2022). Big Data in Food: Systematic Literature Review and Future Directions. *Journal of Computer Information Systems*, 1–21. doi:10.1080/08874417.2022.2132428

Chatterjee, S., Rana, N., & Dwivedi, Y. (2021). How does business analytics contribute to organisational performance and business value? A resource-based view. *Information Technology & People*. Advance online publication. doi:10.1108/ITP-08-2020-0603

Chen, H., Chiang, R., & Storey, V. (2012). Business intelligence and analytics: From big data to big impact. *Management Information Systems Quarterly*, *36*(4), 1165–1188. doi:10.2307/41703503

Cunningham, S., McMillan, M., O'Rourke, S., & Schweikert, E. (2018). *Cracking down on government fraud with data analytics*. Retrieved from McKinsey & Company: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/cracking-down-on-government-fraud-with-data-analytics

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, *17*(4), 263–282. doi:10.1111/j.1467-8551.2006.00500.x

Dwivedi, Y., Janssen, M., Slade, E., Rana, N., Weerakkody, V., Millard, J., Hidders, J., & Snijders, D. (2017). Driving innovation through big open linked data (BOLD): Exploring antecedents using interpretive structural modelling. *Information Systems Frontiers*, *19*(2), 197–212. doi:10.1007/s10796-016-9675-5

EY. (2021). *Big data, big outcomes: how analytics can transform public services and improve citizens' lives*. EYGM Limited.

Ferraris, M. A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision*, *57*(8), 1923–1936. doi:10.1108/MD-07-2018-0825

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *JMR, Journal of Marketing Research*, *18*(1), 39–50. doi:10.1177/002224378101800104

Gartner. (2019). *Gartner Survey Finds Government CIOs to Focus Technology Investments on Data Analytics and Cybersecurity in 2019*. Retrieved from Gartner: https://www.gartner.com/en/newsroom/press-releases/2019-01-23-gartner-survey-finds-government-cios-to-focus-technol

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*, *27*(1), 101–113. doi:10.1016/j.jsis.2017.10.001

Giauque, D., Anderfuhren-Biget, S., & Varone, F. (2013). HRM Practices, Intrinsic Motivators, and Organizational Performance in the Public Sector. *Public Personnel Management*, *42*(2), 123–150. doi:10.1177/0091026013487121

Gopal, P., Rana, N., Krishna, T., & Ramkumar, M. (2022). Impact of big data analytics on supply chain performance: An analysis of influencing factors. *Annals of Operations Research*. Advance online publication. doi:10.1007/s10479-022-04749-6

Gupta, M., & George, J. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049–1064. doi:10.1016/j.im.2016.07.004

Hair, J., Risher, J., Sarstedt, M., & Ringle, C. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. doi:10.1108/EBR-11-2018-0203

Henseler, J., Hubona, G., & Ray, P. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems*, *116*(1), 2–20. doi:10.1108/IMDS-09-2015-0382

Horng, J.-S., Lio, C.-H., Chou, S.-F., Yu, T.-Y., & Hu, D.-C. (2022). Role of big data capabilities in enhancing competitive advantage and performance in the hospitality sector: Knowledge-based dynamic capabilities view. *Journal of Hospitality and Tourism Management*, *51*, 22–38. doi:10.1016/j.jhtm.2022.02.026

Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-Data Applications in the Government Sector. *Communications of the ACM*, *57*(3), 78–85. doi:10.1145/2500873

Kim, S. (2005). Individual-Level Factors and Organizational Performance in Government Organizations. *Journal of Public Administration: Research and Theory*, *15*(2), 245–261. doi:10.1093/jopart/mui013

Klievink, B. R., Cunningham, S., & de Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers*, *19*(2), 267–283. doi:10.1007/s10796-016-9686-2

Ko, J., Hur, S., & Smith-Walker, A. (2013). Family-friendly work practices and job satisfaction and organizational per-formance: Moderating effects of managerial support and performance-oriented management. *Public Personnel Management*, *42*(4), 545–565. doi:10.1177/0091026013505503

Kristoffersen, E., Mikalef, P., Blomsma, F., & Li, J. (2021). Towards a business analytics capability for the circular economy. *Technological Forecasting and Social Change*, *171*, 171. doi:10.1016/j.techfore.2021.120957

Lozada, N., Arias-Perez, J., & Perdomo-Charry, G. (2019). Big data analytics capability and co-innovation: An empirical study. *Heliyon*, *5*(10), e02541. doi:10.1016/j.heliyon.2019.e02541 PMID:31667393

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68. PMID:23074865

Merhi, M., & Bregu, K. (2020). Effective and efficient usage of big data analytics in public sector. *Transforming Government: People*, *Process, and Policy*, *14*(4), 1750–6166.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management*, *30*(2), 272–298. doi:10.1111/1467-8551.12343

Mikalef, P., Krogstie, J., Pappas, I., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, *57*(2), 103169. doi:10.1016/j.im.2019.05.004

Milakovich, M. E. (2012). *Anticipatory Government: Integrating Big Data for Smaller Government. In Internet, Politics, Policy 2012: Big Data, Big Challenges*. Oxford Internet Institute, University of Oxford.

Morabito, V. (2015). Big Data and Analytics for Government Innovation. In V. Morabito (Ed.), *Big Data and Analytics: Strategic and Organizational Impacts* (pp. 23–45). Springer. doi:10.1007/978-3-319-10665-6_2

Munné, R. (2016). Big Data in the Public Sector. In J. Cavanillas, E. Curry, & W. Wahlster (Eds.), *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe* (pp. 195–208). Springer. doi:10.1007/978-3-319-21569-3_11

Murreddu, F., Schmeling, J., & Kanellou, E. (2020). Research challenges for the use of big data in policy-making. *Transforming Government: People*, *Process and Policy*, *14*(4), 593–604.

Pencheva, I., Esteve, M., & Mikhaylov, S. (2020). Big Data and AI – A transformational shift for government: So, what next for research? *Public Policy and Administration*, *35*(1), 24–44. doi:10.1177/0952076718780537

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*(1), 539–569. doi:10.1146/annurev-psych-120710-100452 PMID:21838546

Rainey, H., & Steinbauer, P. (1999). Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations. *Journal of Public Administration: Research and Theory*, *9*(1), 1–32. doi:10.1093/oxfordjournals.jpart.a024401

Rana, N., Chatterjee, S., Dwivedi, Y., & Akter, S. (2022). Understanding dark side of artificial intelligence (AI) integrated business analytics: Assessing firm's operational inefficiency and competitiveness. *European Journal of Information Systems*, *31*(3), 364–387. doi:10.1080/0960085X.2021.1955628

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3*. Boenningstedt: SmartPLS GmbH. Retrieved from http://www.smartpls.com

Rogge, N., Agasisti, T., & De Witte, K. (2017). Big data and the measurement of public organizations' performance and efficiency: The state-of-the-art. *Public Policy and Administration*, *32*(4), 263–281. doi:10.1177/0952076716687355

Santoro, G., Fiano, F., Bertoldi, B., & Ciampi, F. (2019). Big data for business management in the retail industry. *Management Decision*, *57*(8), 1980–1992. doi:10.1108/MD-07-2018-0829

Shmueli, G., Ray, S., Velasquez Estrada, J., & Shatla, S. (2016). The elephant in the room: Evaluating the predictive performance of PLS models. *Journal of Business Research*, *69*(10), 4552–4564. doi:10.1016/j.jbusres.2016.03.049

Ubaldi, B., Van Ooijen, C., & Welby, B. (2019). *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance*. OECD Publishing.

Wamba, S., Gunasekaran, A., Akter, S., Ren, S., Dubey, R., & Childe, S. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365. doi:10.1016/j.jbusres.2016.08.009

Welby, B. (2019). *The impact of digital government on citizen well-being*. OECD Publishing.

Yu, W., Zhao, G., Liu, Q., & Song, Y. (2021). Role of big data analytics capability in developing integrated hospital supply chains and operational flexibility: An organizational information processing theory perspective. *Technological Forecasting and Social Change*, *163*, 163. doi:10.1016/j.techfore.2020.120417

Zhao, J., Fan, S., & Hu, D. (2014). Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics*, *1*(3), 169–174. doi:10.1080/23270012.2014.968643

APPENDIX A

Table 6.
Operationalization of constructs

Construct	Measurement Item	
Basic resources	BR1	We adequately resource our 'big data analytics' projects
	BR2	We are given enough time in our 'big data analytics' projects to achieve our objectives
Data	D1	We integrate data from multiple internal sources into a data warehouse or mart for easy access
	D2	We integrate external data with internal data to facilitate high-value analysis of our business environment
Technology	T1	We have explored or adopted parallel computing approaches (e.g., Hadoop) for big data processing
	T2	We have explored or adopted different data visualization tools
	T3	We have explored or adopted new forms of databases, such as Not Only SQL(NoSQL), for storing data
Technical skills	TS1	We are able to work with advanced statistics (e.g., inferential view of data, linear regression, decision trees)
	TS2	We can research and select the most appropriate tool for our analysis needs
Management skills	MS1	Management perceive data as a source of security and see it as an enabler for progress and support for existing and planned activities
	MS2	Higher management and leaders support data initiatives
Data driven culture	DD1	We base our decisions on data rather than on instinct
	DD2	We are willing to override our own intuition when data contradict our viewpoints
	DD3	We continuously coach our employees to make decisions based on data
Organizational learning	OL1	We are able to acquire new and relevant knowledge
	OL2	We have made concerted efforts to exploit existing competencies and explore new knowledge
Efficiency	EC1	We have made good use of our knowledge and skills in looking for ways to become more efficient
	EC2	We are trying to reduce organizational management and performance costs
	EC3	We conduct business relations with outside customers very promptly
	EC4	We rarely make big mistakes when conducting work
Effectiveness	EV1	In the past two years we have established a significant improvement in productivity
	EV2	Overall, the quality of our work is high
	EV3	Our work provides the public with a worthwhile return on their taxes
	EV4	Achieving goals is very important to us
Fairness	F1	We provide fair and equitable treatment for employees and applicants in all aspects of personnel management without regard to their political affiliation, sex, nationality, marital status, age, or disability
	F2	In general, we treat everyone in our organization with respect, regardless of status and pay grade
	F3	We provide fair and equitable services to the public, regardless of their individual backgrounds
	F4	Our customer satisfaction is very high

APPENDIX B

Table 7.
Cross-loadings

	BR	D	T	TS	MS	DD	OL	EV	EC	F
BR1	**0.948**	0.552	0.546	0.453	0.504	0.532	0.463	0.313	0.335	0.020
BR2	**0.939**	0.535	0.536	0.486	0.428	0.590	0.482	0.358	0.341	0.017
D1	0.343	**0.760**	0.460	0.548	0.350	0.317	0.364	0.345	0.286	0.149
D2	0.594	**0.982**	0.592	0.632	0.554	0.473	0.522	0.346	0.350	0.126
T1	0.524	0.339	**0.772**	0.448	0.352	0.378	0.218	0.219	0.255	-0.087
T2	0.440	0.636	**0.856**	0.565	0.461	0.378	0.367	0.329	0.360	0.107
T3	0.452	0.427	**0.827**	0.562	0.397	0.459	0.331	0.263	0.306	0.021
TS1	0.538	0.655	0.600	**0.959**	0.459	0.525	0.636	0.428	0.524	0.176
TS2	0.415	0.607	0.624	**0.959**	0.476	0.485	0.539	0.490	0.552	0.159
MS1	0.376	0.438	0.419	0.405	**0.871**	0.459	0.427	0.459	0.424	0.299
MS2	0.491	0.518	0.460	0.451	**0.887**	0.517	0.598	0.445	0.451	0.347
DD1	0.583	0.440	0.437	0.469	0.551	**0.906**	0.554	0.484	0.408	0.170
DD2	0.450	0.272	0.317	0.380	0.336	**0.807**	0.425	0.346	0.310	0.089
DD3	0.495	0.486	0.470	0.505	0.532	**0.867**	0.555	0.531	0.440	0.219
OL1	0.431	0.493	0.350	0.571	0.552	0.542	**0.930**	0.589	0.590	0.454
OL2	0.500	0.479	0.352	0.571	0.541	0.571	**0.934**	0.494	0.567	0.343
EV1	0.146	0.195	0.221	0.180	0.353	0.216	0.244	**0.671**	0.484	0.554
EV2	0.238	0.240	0.196	0.313	0.451	0.361	0.453	**0.793**	0.473	0.630
EV3	0.397	0.336	0.373	0.406	0.381	0.538	0.458	**0.818**	0.664	0.286
EV4	0.257	0.343	0.235	0.509	0.430	0.447	0.583	**0.827**	0.601	0.534
EC1	0.090	0.140	0.109	0.193	0.283	0.158	0.365	0.440	**0.552**	0.503
EC2	0.168	0.222	0.184	0.355	0.303	0.150	0.427	0.520	**0.608**	0.462
EC3	0.403	0.399	0.386	0.591	0.464	0.487	0.579	0.655	**0.897**	0.334
EC4	0.262	0.210	0.333	0.379	0.371	0.381	0.418	0.496	**0.800**	0.283
F1	-0.012	0.021	0.044	0.063	0.282	0.099	0.256	0.486	0.358	**0.815**
F2	0.024	0.154	0.026	0.219	0.333	0.270	0.385	0.577	0.414	**0.800**
F3	-0.024	0.080	-0.007	0.110	0.261	0.040	0.286	0.429	0.345	**0.787**
F4	0.045	0.134	0.027	0.099	0.271	0.087	0.370	0.388	0.393	**0.800**

APPENDIX C

Table 8.
Heterotrait-Monotrait ratio of correlations (HTMT)

		4	5	6	7	8	9	10
4	Technical skills							
5	Management skills	0.607						
6	Data driven culture	0.605	0.720					
7	Organizational learning	0.696	0.754	0.709				
8	Effectiveness	0.529	0.692	0.611	0.677			
9	Efficiency	0.645	0.686	0.521	0.788	0.956		
10	Fairness	0.174	0.461	0.182	0.478	0.761	0.690	

Guido Ongena is an associate professor and program director the HU University of Applied Sciences Utrecht. He has an MSc in Business Informatics from the Utrecht University (2007) and a PhD from the University of Twente (2013). His research interests include quantitative research, technology adoption and social network analysis. He is specifically focused on the application of big data analytics. His research is published in several international journals and conferences, including Business Process Management Journal, Behaviour and Information Technology, Information Research, and Telematics and Informatics.

Arjen Davids is a data analyst at the municipality of Berg en Dal.