Insights from the CGMA Data Competencies Model: The Role of Data Culture to the Value Creation Process

Wong Kee-Luen, Julian Teh Hong-Leong, Ng Shwu-Shing, Chuah Min-Hooi
Universiti Tunku Abdul Rahman, Kampar Campus, Malaysia.
wongkl@utar.edu.my

***Abstract*—With the emergence of the big data phenomena, the business intelligence maturity approach tends to be limiting and lacks the capability to capture and engage with the relevant variables and develop into a theoretical framework to explain the big data economy. The concept of data competencies proposed by Chartered Global Management Accountants (CGMA) was thought to be a more comprehensive alternative framework to explore the phenomena. The four types of data competencies, namely, data culture, data management, data analytics and value creation were used to construct the conceptual framework to understand and explain the big data initiative implementation process. It was found that data culture tends to moderate the data management-data analytics relationship. In addition, data analytics appears to partially mediate the impact of data management on value creation. The implications of these findings confirm that data culture is the essential foundation to the value creation process.**

***Index Terms*—Business Intelligence; Big Data; Data Competencies; Value Creation.**

I. Introduction

Business intelligence (BI) experts tend to focus on BI systems as tools that enable them to find and get information from data sources [1][2]. Many authors in the field of IS make use of maturity models to benchmark and assess the competence of an organization to implement BI system successfully [3][4][5]. For example, Gartner's maturity model can be used to rate business maturity levels and the maturity of respective departments [6]. The proposed five maturity levels: unaware, tactical, focused, strategic, and pervasive. Other maturity models including TDWI's maturity model, Hewlett Packard's business intelligence maturity model, Business information maturity model, AMR Research's BI/performance management maturity model, Business intelligence development model [7]. These maturity models together with our Enterprise Business Intelligence Maturity Model (EBIMM) included tending to be descriptive rather than predictive theoretical models [7].

Descriptive models such as business intelligence maturity models tend to evaluate the indicative maturity status of the IT systems of business organizations. Theoretical models, however, tend not only to explain the investigated phenomena in terms of the relationships between variables and constructs, but also predict the becoming of the dependent variable in the vents of changes to the independent variables [8].

Thus, to further the understanding of the big data phenomenon, a theoretical framework is required that is capable of capturing the relevant variables (both dependent and independent variables) surrounding the big data economy.

The Chartered Global Management Accountants (CGMA) Report pointed out that the priority of business organizations is to data mine the readily available streams of data in their IT systems [9]. The imminent weakness among the business organizations is the lack of skills and competencies to capture the promising opportunities and benefits of the big data phenomenon [10]. In addition, CGMA presented the big data competencies model and proposed that business organizations will require new abilities and competencies: data culture, data management, data analytics, and value creation in order to capture and realize the opportunities and benefits of the big data economy [9].

The CGMA data competencies model was found to be attractive and the plausibility of developing a theoretical framework based on data competencies. Therefore, based on CGMA data competencies model, the conceptualized variables are data culture, data management, data analytics and value creation. The interconnections between data culture, data management, data analytics and value creation could also be explored.

II. Literature Review

Viewed from the accounting perspective, the big data phenomenon is a relatively new concept. The Malaysian Institute of Accountants (MIA) only reported the phenomenon in the November/December 2014 issue. Thus, there is relatively little literature related to the issue.

The Chartered Global Management Accountants (CGMA), however, spearhead to examine the phenomenon by starting the CGMA Briefing on Big Data with the purpose of readying business for the big data revolution [9]. In addition, they presented the big data competencies model. The required competencies range from technical ability to business acumen and span from performance management to conformance to data management standards (see Figure 1).

Based on the model, business organizations required the following abilities and competencies:

- Data culture – the culture that decisions are made objectively and based on analysis of available data and evidence.
- Data management – businesses need to ensure that their IT systems ensure data integrity, that data are captured correctly and relevantly, that data stored are accessible for consistent use.

- Data analytics – advanced level of analytical skills for data mining, deriving algorithms, and predictive analytics
- Value creation – the ability to translate analytical insights into commercial insights, and business acumen to identify opportunities.

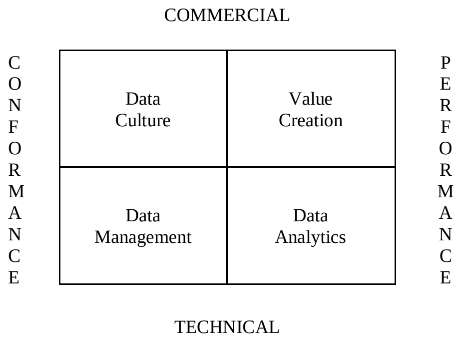

Figure 1: The range of big data competencies
(adapted from CGMA, 2014 [9])

Chuah and Wong presented thirteen competency areas related to information systems initiatives and implementation. They are change management, people management, culture, knowledge management, infrastructure, data warehousing, master data management, meta data management, analytics, performance management, balanced scorecard, information quality, and strategic management [11].

These thirteen BI competency areas were assigned to the four data competencies accordingly, such that:

- Data culture – People, Organizational culture, Change management
- Data management – Data warehouse, Master Data management, Infrastructure, Information quality
- Data analytics – Metadata management, Knowledge management, Analytical
- Value creation – Performance management, Balanced scorecard, Strategic management

III. THEORETICAL FRAMEWORK DEVELOPMENT

To develop the theoretical framework that can be used to understand the big data phenomenon better, value creation was identified as the dependent variable. This is because from the management accounting perspective, value creation is the ultimate objective and goal of the big data economy. The subsequent independent variables used to explain value creation are data culture, data management, and data analytics.

Organizations are in their various degrees of data competencies while attempting to capture the benefits of big data. Generally, "a business needs to have the right data, the ability to analyze it … and the skills to ensure that insight are applied to create value" for the business [9]. Thus, we theorize that an organization's strategy to achieve competitive advantage from big data starts from data management when their IT "systems and processes capture relevant data correctly, first time, and store it accessibly for consistent use" [9]. The next stage in the process would be data analytics where business organizations have the advanced level of analytical skills including data mining, algorithms and predictive analytics to generate reports and analyses that can be subsequently translated from analytical insights to commercial opportunities to create business value – value creation, which is the ultimate goal of tapping into big data. To ensure the success of the three stages: Data management, Data analytics, and Value creation, top management support in the form of data culture are necessary whereby "data are valued as an important strategic asset" and "decisions are based on analysis of available data and evidence" [9].

To construct an adequate theoretical model relating to the big data phenomenon, "one must probe the mechanisms that underlie an effect and the limiting conditions for its occurrence" [12].

Understanding the mechanisms of the effect produces more refined assessments of what the effect really is and how it is produced while understanding its limiting conditions inform the study about the necessary conditions for the effect to occur [12]. In addition, these two types of understanding – one of mechanisms and one of limiting conditions, are the concerns of mediation analyses and moderation analyses respectively. The fact is "the understanding of mechanisms and the understanding of limiting conditions are theoretically intertwined, and in combination, give rise to a full theoretical understanding of the effect of interest" [12].

Following Judd et al., the theoretical framework was re-modeled for the study. Sekaran and Bougie commented that experience and intuition play an important role in theoretical framework development [8]. The big data initiative process starts with data management where relevant and good quality data are captured, stored and could be readily retrieved for use. Subsequently these data have to be analyzed to generate reports and analytical results via the data analytics stage. These reports and analyses have to be interpreted and translated into business insights before value can be created for the organization. Thus, the big data initiative process is a three-stage process, with data management having direct effect on value creation via data analytics. In other words, value creation comes about due to the mediating effects (the mechanisms) of data analytics on the data management-value creation relationship. To state the framework clearly, data management as an independent variable, acting alone, would have lesser influence on value creation, but data management acting with data analytics would have a much stronger impact on value creation.

The other question is: What role does data culture play in this big data initiative process? It is commonly known that no big data initiatives would be successful without top management support and commitment to substantial investments in big data initiatives [9]. Thus, data culture is the prerequisite foundation of all big data initiatives where strategic decisions are made based on available data and evidence.

Therefore, we propose that data culture is the moderating variable or the necessary and sufficient conditions for successful data analytics and subsequently value creation to happen.

The theoretical framework is shown in Figure 2.

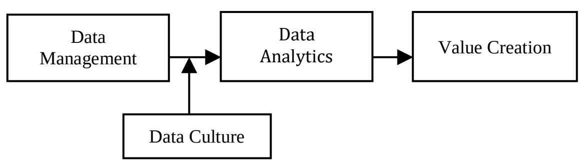

Figure 2: CGMA's big data initiative process theoretical framework

Based on the theoretical framework shown in Figure 2, the following two hypotheses are proposed.

H1: Data culture moderates the relationship between Data management and Data analytics.

H2: Data analytics mediates the impact of Data management on Value creation.

IV. Research Methodology

Data were collected from 132 business organizations from various sectors such as construction, financial/banking, manufacturing, using the questionnaire developed by Chuah, et al. [11]. The questionnaire captured data on fifty-two (52) items based on 5-point Likert scale. These fifty-two (52) items were factored into the thirteen competency areas or constructs which were further assigned to the four data competencies or variables.

Data management, the independent variable was assigned four constructs – Data warehouse, Master data management, Infrastructure, and Information quality. Data culture, the moderating variable was assigned three constructs – People, Organizational culture, and Change management. Data analytics, the mediating variable was assigned three constructs – Metadata management, Knowledge management, and Analytical. And Value creation, the dependent variable was assigned three constructs – Performance management, Balanced scorecard, and Strategic management. Table 1 shows the variables, constructs and the number of measuring items.

Table 1
Schedule of variables, constructs, and items

Variables	Constructs	No. of items
Data management	Data warehouse	4
	Master data management	7
	Infrastructure	1
	Information quality	4
Data culture	People	5
	Organizational culture	3
	Change management	2
Data analytics	Performance management	3
	Balanced scorecard	4
	Strategic management	4
Value creation	Performance management	5
	Balanced scorecard	4
	Strategic management	6

To test the moderating effects of data culture, we made use of the multiple linear regression (MLR) method. Following Dawson, we assigned Data analytics (DA) as the dependent variable and Data management (DM) as the independent variable and Data culture (DC) as the moderating variable [13]. We created a new interaction variable (DM*DC) and performed two regression analyses using Statistical Packages for Social Sciences (SPSS) Version 20. The first regression was to regress DM and DC on DA. The second regression was to regress DM, DC and DM*DC on DA. The results could be interpreted to inform about the moderating effect.

To test for the mediating effect of DA on the impact DM on value creation (VC), we followed Andrews, Goes and Gupta, who suggested that four specific criteria must be met: (1) the independent variable (DM) should significantly influence the mediator (DA); (2) the mediator (DA) should significantly influence the dependent variable (DV); (3) the independent variable (DM) should significantly influence the dependent variable (VC); (4) after the mediator variable (DA) is controlled for, the impact of the independent variable (DM) should no longer be significant (for full mediation) or should be reduced (for partial mediation) [14]. We used a partial least square (PLS) structural equation modeling (SEM) technique to identify the mediating effect.

Table 2
Results of measurement properties

Variable name	Cronbach alpha	Composite reliability (CR)	AVE
Data culture	0.882	0.927	0.809
Data management	0.772	0.866	0.683
Data analytics	0.702	0.870	0.770
Value creation	0.896	0.935	0.827

Table 3
Correlations

Variable name	(1)	(2)	(3)	(4)
Data culture (1)	0.899			
Data management (2)	0.873	0.827		
Data analytics (3)	0.589	0.716	0.877	
Value creation (4)	0.846	0.884	0.798	0.910

Note: Figures in diagonal are values of the square root of the AVE

V. DATA ANALYSIS AND RESULTS

The variables were checked for reliability, convergent validity and discriminant validity. Table 2 shows the Cronbach alpha values, composite reliability (CR), and the average variance extracted (AVE).

The Cronbach alpha values ranged from 0.702 to 0.896 for the four variables. All the Cronbach alpha values exceed the 0.70 threshold [15], indicating high internal reliability. Similarly, all composite reliabilities (CR) were also high and ranged from 0.866 to 0.935 (see Table 3) indicating high reliability. Therefore, internal reliabilities of the variables were confirmed. Convergent validity was assessed by reviewing the indicator loadings. All indicator loadings for each variable were significant. The average variance extracted (AVE) values ranged from 0.683 to 0.827, meaning that all the AVE values were above the recommended threshold of 0.50 [16], proving convergent validity for all the variables.

The discriminant validity of the variables was assessed by examining the correlations of the variables. The values of the square root of the AVE (shown in diagonal in Table 3) were all greater than the off-diagonal correlations. Therefore, the discriminant validity was confirmed [17].

We proceeded to subject our two hypotheses presented in Section 3 to empirical testing.

Hypothesis 1: *Data culture moderates the relationship between Data management and Data analytics.*

Following Chin, Marcolin, and Newsted, we performed two regression analyses with SPSS Version 20 [18].

Model 1:

$$DANALYTICS = C + \beta_1 DCULTURE + \beta_2 DMANAGEMENT + \varepsilon$$

Model 2:

$$DANALYTICS = C + \beta_1 DCULTURE + \beta_2 DMANAGEMENT + \beta_3 DCULTURE*DMANAGEMENT + \varepsilon$$

The results of the regression analyses are shown in Table 4 below:

Table 4
Results of Regression Models

	R	R^2	B	β	t	Sig.
Model 1	0.797	0.635				
(Constant)			2.968	-	78.904	0.000
DCULTURE			0.068	0.070	0.626	0.532
DMANAGEMENT			0.874	0.735	6.603	0.000
Model 2	0.847	0.718				
(Constant)			2.734	-	51.706	0.000
DCULTURE			0.539	0.553	4.398	0.000
DMANAGEMENT			0.923	0.776	7.888	0.000
DCULTURE*DMANAGEMENT			0.656	0.595	6.146	0.000

Dependent variable: DANALYTICS

Table 5
Results of PLS for mediation effects

	Model 1 (IV for MV)	Model 2 (MV for DV)	Model 3 (IV for DV)	Model 4 (Control for DV)
Data management →Data analytics	0.814**	-	-	0.761**
Data analytics →Value creation	-	0.803**	-	0.264**
Data management →Value creation	-	-	0.885**	0.699**
R^2				
Data analytics	0.663	-	-	-
Value creation	-	0.645	0.784	0.839

** Significance at 0.01

Table 4 shows DCULTURE*DMANAGEMENT (β = 0.595, p = 0.000). In addition, the results also give a standardized coefficient (β) of 0.553 from DCULTURE, 0.776 from DMANAGEMENT with R-square of 0.718. These results imply that one standard deviation increase in DCULTURE will impact DANALYTICS by 0.553, but it would also increase the impact of DMANAGEMENT on DANALYTICS. The main effects (see Model 1) as expected resulted in a slightly lower standardized beta (β = 0.735) and a smaller R-square of 0.635. The interaction effect has a calculated effect size of 0.294 which lie between medium and large effect [18]. The results confirmed the interaction effect and therefore Hypothesis 1 is supported.

Hypothesis 2: *Data analytics mediates the impact of Data management on Value creation.*

Following Chen and Tsou, we adopted Andrew et al.'s four criteria for establishing mediating effect [19]. Table 5 shows the results of the mediating effects.

We tested the four conditions for establishing mediating effects using PLS-SEM analysis. Table 5 shows Model 1 and Model 2 meeting the first and second criteria. This means Data management (IV) has significant influence on Data analytics (MV) (β=0.814, p<0.01). Similarly, Data analytics (MV) has significant influence on Value creation (DV) (β=0.803, p<0.01). Model 3 satisfies the third criteria, that is, Data management (IV) has significant impact on Value creation (DV) (β=0.885, p<0.01). Model 4 results show that including Data analytics as the mediator decreases the impact of Data management (IV) on Value creation (DV) (β=0.699, p<0.01). Although the impact of Data management on Value creation decreased, from 0.885 to 0.699, the influence remains significant indicating that Data analytics exerts partial mediating effect on Value creation. Therefore, Hypothesis 2 is supported.

VI. DISCUSSIONS

The main objective of this study is to probe and understand the underlying mechanisms of the mediating effects of data analytics to value creation, and also the necessary and sufficient conditions for value creation to be produced. For any big data initiative to succeed, top management support is of utmost importance. In addition, the people and culture of the organization should facilitate and participate in the big data initiative undertaken. Continuous investments in data management (data warehouse, master data management, information quality) and data analytics (metadata management, knowledge management, analytical systems) are inevitable and should not be seen as wasteful. The pre-requisite culture to inculcate is that strategic and managerial decisions are made objectively and are based on analysis of available data and evidence [9].

Data culture (people, organizational culture, change management) is the necessary and required conditions whereby data analytics can become effective to generate insights to value creation. Without a supportive culture for big data initiatives, and with data management acting alone, data analytics will most definitely be less effective to create value for the company.

Our results have indeed suggested that data analytics has partial mediating effects on the influence of data management on value creation. This means that data analytics is needed to improve the effectiveness of value creation. This mediating effect explains the difference between the direct effect from the direct residual effect. In our case, the direct effect is 0.885 (see Table 6, Model 3) and the direct residual effect is 0.699 (see Table 6, Model 4). There is also evidence of suppression in the mediation model. Suppression occurs when there is significant indirect effect (a*b) and significant direct residual effect (c'), among which the sum of the effects is greater than the original direct effect (c) (Judd, et al. 2012). In our case, the indirect effect is 0.201 (0.761 x 0.264) and the direct residual effect is 0.699 and their sum of 0.900 is greater than direct effect (0.885). Thus, there is evidence of suppression in the model. When suppression occurs, the mediator tends to dampen the direct effect. So the inclusion of data analytics in the mediation model leads to dampening of the total direct effect of 0.885 to 0.699 (direct residual effect).

The mediation analysis was conducted in order to understand the mechanism that "produces more refined assessments of what the effect really is and how it is produced" [11]. In our case, we proved that data analytics is the partial mediator and is responsible for the data management-value creation relationship. (See Table 6, Model 4.) In other words, data analytics mediates the influence of data management on value creation. In Model 4, the indirect effect is 0.201 (0.761 x 0.264) and the residual direct effect is 0.699 with R^2 of 0.839. Compared to Model 3 where the direct effect is 0.885 with R^2 of 0.784. The total effect of Model 4 is 0.900 (0.201 + 0.699) which is slightly greater than the direct effect of Model 3.

By including data analytics as the mediator, the variance explained improved from 78.4% to 83.9% and the total effect on value creation also increased by 0.015, from 0.885 to 0.900. Thus, to improve the mediating effects of data analytics, continuous increased investments on advanced analytical systems and knowledge management systems are desired. These investments when managed professionally should be able to generate analytical insights from information retrieved from data management. Additionally, advanced level of analytical skills such as algorithms and predictive analytics are required. This involves hiring personnel with advanced data skills to staff the data analytics process. Further, qualified personnel are needed to translate analytical insights into commercial insights so as to create value. Apart from new business opportunities to generate extra revenue, value can be created by increasing efficiency, reducing risk, and improving cash flow.

Integrating the moderating effect of data culture on the data management-data analytics relationship, and the mediating effects of data analytics on the influence of data management on value creation, we propose to re-arrange the variables of the CGMA Data Competencies Model to better reflect the underlying mechanisms of the mediating effects of data analytics to value creation, and also the necessary and sufficient conditions for value creation be produced. The new amended model is presented below:

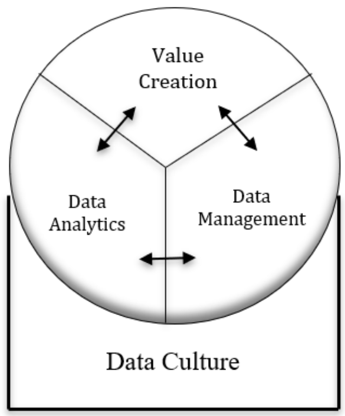

Figure 2: The modified CGMA Data Competencies Model (developed for the study)

We found the CGMA data competencies framework to be an adequate model to explain the big data initiative process [14]. Any big data initiative begins with the setting up of data management systems, subsequently advances to data analytics procedures in order to generate analytical insights. These analytical insights are translated into commercial insights so that value can be created for the firm. The value creation process proceeds from data management to data analytics and then from data analytics to value creation. Feedbacks on the limitations of the analytical insights from the value creation stage then reverse backward to the data analytics stage and subsequently to data management stage to request for timely information data. The value creation process therefore would proceed back and forth from the data management stage to the data analytics stage, and to value creation stage.

Due to data analytics only exhibiting partial mediating effects between data management and value creation, we postulate that there should still be a significant link between data management and value creation. At times, the data management stage could go directly to the value creation stage and the value creation stage could request for more relevant information for decision making. In other words, it is possible that value-creating insights could be extracted straight from data management thus by-passing data analytics.

All the three stages should be supported by data culture whereby data are valued as strategic assets and decisions are made based on evidence and valid data analysis. The supportive role of data culture to the value creation process cannot be looked upon lightly. In other words, data culture acts as a cradle to the value creation process. Data culture is the foundation of the value creation process and has to be strong in order for insightful and creative ideas to be generated. Managers are expected to preview and digest available information and data for decision making as part of their daily tasks. The IT division would have to work with human resources to facilitate this cultural transformation.

VII. Conclusion

The study set out to look for an alternative conceptual model (instead of the enterprise business intelligence maturity model), to capture the relevant variables sufficiently about the emerging big data phenomena.

We used data that were collected by the survey questionnaire and found that all the variables (data culture, data management, data analytics, and value creation) are statistically reliable and valid – both convergent validity and discriminant validity could be established. In addition, we structured the conceptual framework in such a manner that data management was the independent variable, data culture – moderating variable, data analytics – the mediator, and value creation – the dependent variable. Using multiple regression analyses (MLA), we proved that data culture exhibited interaction effects on the data management-data analytics relationship. We also managed to establish that data analytics mediates the impact of data management on value creation, using PLS-SEM technique.

We achieved the objective of this study, using data collected for our study based on thirteen competency areas. It is recommended that a new proprietary questionnaire or measuring instrument be designed specifically to collect data based on the CGMA data competencies, and also over wider geographical areas. In addition, bigger samples should be selected to ensure better representation of the target population with the benefit of higher external validity. Until then we can only accept the results and findings with caution and due care.

References

[1] Damjanovic, V. & Behrendt, W. "UNDERSTANDER: Business Intelligence Seeker – User Agent", *37th Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1491 – 1496, 2014.

[2] Yoon, T. S., Ghosh, B. & Jeong, B. K.. "User Acceptance of Business Intelligence (BI) Application: Technology, Individual Difference, Social Influence, and Situational Constraints". *47th Hawaii International Conference on System Sciences (HICSS)*, pp. 3758 – 3766, 2014.

[3] Harpham, A. *The APM Group's assessment model for portfolio, program and project management, its PRINCE2 maturity model and their benefits to organizations*, Available from :<http://www. apmgroup co uk/nmsruntime/ saveasdialogasp?lID=576&sID=102>. [Retrieved: 27 December 2009].

[4] Paulk, M. C., Curtis,B., Chrissis, M. B. & Weber, C. "Capability Maturity Model for Software, Version 12", *Software Engineering Institute/Carnegie Mellon University*, 2006.

[5] Rajterič, I. H. "Overview of Business Intelligence Maturity Models", *International Journal of Human Science*. Vol. 15, No. 1, pp 47-67, 2010.

[6] Gartner Research, *IT Score Overview for Business Intelligence and Performance Management.* Available from :<http://www gartner com/resources/205000/205072/ itscore_overview_for busines_205072 pdf>. [Retrieved: 11 November 2010].

[7] Wong, K. L., Chuah, M. H. & Ong S. F. "Are Malaysian companies ready for the big data economy? A business intelligence approach", *.International Conference on Accounting Studies (ICAS) 2015*, 17-20 August 2015, Johor Bahru, Johor, Malaysia.

[8] Sekaran, U. & Bougie, R. *Research Methods for Business: A Skill Building Approach*, 5th Ed. John Wiley & Sons: West Sussex, UK, 2009.

[9] CGMA Report, *CGMA briefing - Big data: Readying business for the big data revolution*, 2014.

[10] McKinsey Report, *Views from the front lines of the data revolution*, McKinsey & Co, 2014.

[11] Chuah, M. H and Wong, K. L., "A framework for accessing an Enterprise Business Intelligence Maturity Model: Delphi study approach", *African Journal of Business Management*, Vol.6 (23), pp 6880-6889, 2012.

[12] Judd, C. M., Yzerbyt, V. Y., & Muller, D., "Mediation and moderation", *Handbook of research methods in social and personality psychology*, Vol. 2, pp. 653-676, 2014.

[13] Dawson, J. F., "Moderation in management research: What, why, when, and how", *Journal of Business Psychology*, 29: 1-19, 2014.

[14] Andrews, J. C., Goes, P. B., & Gupta, A. "Understanding adolescent intentions to smoke: an examination of relationships among social influences, prior trial behaviors, and antitobacco campaign advertising", *Journal of Marketing*, 68,110-123, 2004.

[15] Nunnally, J., *Psychology Theory*. New York: McGraw-Hill, 1978.

[16] Barclay, D. W., Thomson, R., Higgins, C., "The partial least squares (PLS) approach to causal modelling: personal computer adoption and use an illustration", *Technology Studies*, 2 (2), 285-309, 1995.

[17] Fornell, C., & Larcker, D. F., "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, 19, 39-50, 1981.

[18] Chin, W. W., Marcolin, B. L., & Newsted, P. R., "A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and voice mail emotion/adoption study", *17th International Conference on Information Systems*, 16-18 December 1996, Cleveland, Ohio, 1996.

[19] Chen, J. S. & Tsou, H. T., "Performance effects of IT capability, service process innovation, and the mediating role of customer service", *Journal of Engineering and Technology Management*, 29: 71-94, 2012.