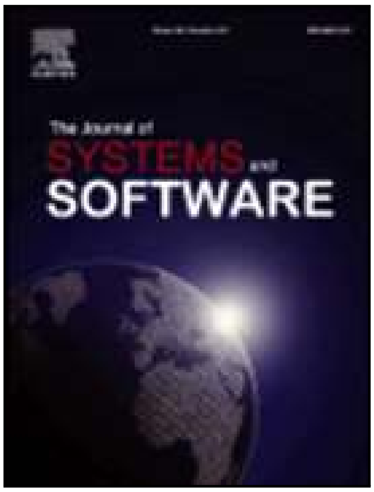

Agile Values or Plan-Driven Aspects: Which Factor Contributes More toward the Success of Data Warehousing, Business Intelligence, and Analytics Project Development?

Dinesh Batra Ph.D. Professor

PII: S0164-1212(18)30212-7
DOI: https://doi.org/10.1016/j.jss.2018.09.081
Reference: JSS 10229

To appear in: *The Journal of Systems & Software*

Received date: 21 November 2017
Revised date: 25 September 2018
Accepted date: 26 September 2018

Please cite this article as: Dinesh Batra Ph.D. Professor , Agile Values or Plan-Driven Aspects: Which Factor Contributes More toward the Success of Data Warehousing, Business Intelligence, and Analytics Project Development?, *The Journal of Systems & Software* (2018), doi: https://doi.org/10.1016/j.jss.2018.09.081

Highlights

- Analytics development can be done in agile-plan balanced or agile-heavy mode
- Agile values are essential for both agile-plan balanced and agile-heavy modes
- Plan-driven aspects are essential only for the balanced mode
- Top management is a crucial antecedent for the agile-plan balanced mode
- Shared understanding is a crucial antecedent for the agile-heavy mode

Agile Values or Plan-Driven Aspects: Which Factor Contributes More toward the Success of Data Warehousing, Business Intelligence, and Analytics Project Development?

Dinesh Batra

Professor, Information Systems and Business Analytics, College of Business Administration, Florida International University, Miami, FL 33199 USA, batra@fiu.edu

Abstract

Practically all organizations are developing data warehousing, business intelligence, and analytics (DW/BIA) projects for achieving customer value. A DW/BIA development project may be characterized by both agile and plan-driven aspects. The reported study investigated two research questions: 1) Which factor, agile values or plan-driven aspects, contributes more toward the success of DW/BIA? 2) What are the significant antecedents of agile values and plan-driven aspects? 124 respondents engaged in DW/BIA development filled a 30-item questionnaire on seven constructs. The partial least squares structural equation modeling (PLS-SEM) method was used to determine the strength of the relationships among the following factors: technological capability, shared understanding, top management commitment, and complexity as antecedents; agile values and plan-driven aspects as mediating; and project success as the dependent construct. Based on a prediction-oriented segmentation (PLS-POS) analysis, the findings indicate that there are two groups, agile-plan balanced and agile-heavy, which represent different approaches to DW/BIA development. Top management commitment and shared understanding emerge as strong antecedents to agile values and plan-driven aspects. Overall, the factor agile values contributes more toward the success of DW/BIA development.

Keywords: Agile values; Plan-Driven; Top Management Commitment; Shared Understanding; Business Intelligence; Analytics

1. Introduction

In their classic book, Boehm and Turner (2004) suggested that the characteristics of plan-driven and the agile approaches for software development need to be harmonized and balanced. The plan-driven methods are based on process management (Boehm & Turner, 2004) whereas the agile methods are based on the values and principles stipulated in the Agile Manifesto declared in 2001 (Fowler & Highsmith, 2001; Nerur, Mahapatra, & Mangalaraj, 2005). Methods such as Scrum (Schwaber, 2004) have been widely used to implement agile values and principles in practice (VersionOne, 2016). Although agile methods were initially reported for small-size, dynamic projects (Boehm & Turner, 2004; Dyba & Dingsoyr, 2008), these methods have been scaled up for larger software projects, and it seems that agile methods are widespread and are incorporating some of the practices of plan-driven approaches (Batra, VanderMeer, & Dutta, 2011; Knaster & Leffingwell, 2017; Larman & Vodde, 2016). Conversely, plan-driven methods are being coordinated with agile practices (Ambler, 2009). A contrast between process-centric and agile methods is provided by Nerur et al. (2005), but it seems that recent developments suggest a harmonization of the two approaches.

It would seem, therefore, that a given software development product/project would display essential elements of both plan-driven and agile approaches. In this paper, these essential elements are termed as aspects. Plan-driven aspects are key process elements based on planning, control, risk management, customer expectation management, and adherence to contracts and promises (Humphrey, 1995; Kan, 2002). Agile aspects emphasize individuals and interactions, working software, customer collaboration, and responding to changes (Fowler & Highsmith, 2001). Because of the clear stipulation of a manifesto, agile aspects may be equated to agile values although one may optionally consider some agile principles and practices, too. Plan-driven and agile aspects are not entirely mutually exclusive; the former represents a top-down approach, and the latter represents a bottom-up approach for achieving similar project/product objectives. Both have a similar goal of achieving project success by improving quality and customer satisfaction and lowering risk. For example, the Rational Unified Process (RUP) is regarded as a plan-driven approach based on iterative development (Kruchten, 2000) and can be harmonized with the agile approach (Ambler, 2009). Similarly, an agile method like Scrum (Schwaber, 2007) has its practices of instituting planning and control (e.g., iteration planning, daily standups) and managing risks (e.g., iteration reviews) although these aspects are not usually conducted in upfront and management-driven manner.

DW/BIA project development seems to exhibit characteristics of both agile and plan-driven aspects. The DW/BIA staged-model may favor plan-oriented stages such as integration of source systems and architecture development whereas the iterative nature of analytics model evaluation may favor the agile aspects; however, these are simplistic assumptions that are not anchored in the recent empirical literature. According to Hughes (2008) and Collier (2012), the data warehousing, business intelligence, and analytics development has traditionally been considered compatible with planning-based approaches because it was falsely assumed that data modeling implementations are rigid and difficult to adjust. Collier (2012) asserts: "The reality is that there is nothing special about data-centric systems that makes Agile principles irrelevant or inappropriate. The challenge is that Agile practices must be adapted…" (P. xxv). However, even as a strong agile proponent, Collier (2012) argues: "It is essential that we have a sufficient amount of planning…we must able to adapt our plan to changing factors" (p. 5). There is the preliminary empirical support that DW/BIA development requires both agile and plan-driven aspects (Batra, 2017). It is not clear, however, whether both agile values and plan-driven aspects are about equally important, whether one is more important than the other, or if neither is essential for project success. By using the survey method, the reported study empirically and quantitatively examines which aspect is more important for project success, and what are the significant antecedents to these aspects.

Business Intelligence and Analytics have emerged as an essential area of study for both practitioners and researchers reflecting the magnitude and impact of data-related problems to be solved in contemporary business organizations (Chen, Chiang, & Storey, 2012). Data warehousing (DW) provides the foundation for this decision support infrastructure (Ariyachandra & Watson, 2010). Business analytics (BA) helps to understand the information contained in the data and derive insights that are most important to future business decisions (Sharda, Delen, & Turban, 2016). Business Intelligence (BI) combines architecture, databases, data warehouses, analytical tools, and applications (Sharda et al., 2016).

DW/BIA projects are sometimes viewed as consisting of stages that reflect its many facets. Before the DW/BIA algorithms can be run to improve prediction, the data set needs to be prepared diligently (Siegel, 2016). According to Collier (2012), there are four broad stages: source systems, integration, presentation, and analysis. The data warehouse architecture includes one or more source systems from which data is extracted, transformed, and loaded (ETL) into the data warehouse repositories. Data from these sources is transported into an integration tier where data can be merged, manipulated, cleansed, and validated without placing an undue burden on the operational systems. Data is then moved into a presentation tier (e.g., star schema or a variant), which is more amenable to optimized multidimensional and analytical queries. Finally, data is presented to the users at the analysis tier by using visualization, data mining, statistical analysis, and other BIA technologies (Abbasi, Sarker, & Chiang, 2016; Chen et al., 2012). Despite the normative stage model of a DW/BIA project, there is no empirical evidence that the definition of stages in the development of a DW/BIA product connotes a preference for plan-driven aspects (Collier, 2012).

The variety of activities involved in DW/BIA development suggests that both agile and plan-driven aspects may be required. It is premature to assume that the development of analytics projects or products is more amenable to either plan-driven or agile aspects. In the contemporary development environment, a holistic approach is essential given that the relative importance of plan-driven aspects versus agile values can only be determined empirically by assessing how practitioners engaged in DW/BIA domain view each factor important for project development success. If agile values and plan-driven aspects are critical for DW/BIA project success, then it is also essential to study the antecedents that foster the two constructs. Based on a qualitative study (Batra, 2017), four factors - shared understanding, technological capability, top management commitment, and complexity - are considered as antecedents for this study.

The research provides a holistic framework to examine the relative roles of agile values and plan-driven aspects for achieving DW/BIA success and identifies the antecedents of agile values and plan-driven aspects. The paper provides scales used for the survey questionnaire thereby allowing other researchers to replicate or conduct similar research. The next sections report the research instrument design, methodology, results, discussion, limitations, and suggestions for future research.

2. Research Instrument Design from Literature

Based on interviews with senior managers, a recent qualitative study examined agile practices for DW/BIA project development and found that both agile development and project management aspects seemed to affect business value (Batra, 2017). Furthermore, this qualitative study identified five factors – technological capability, shared understanding, top management commitment, complexity, and organizational culture – as possible antecedents of agile development and project management. Because of the typical limitations of qualitative research, the (Batra, 2017) study could not establish statistical significance or determine the magnitude of the effects.

Plan-driven and agile aspects are distinct constructs that may be present in different amounts in a project. The two aspects also have some overlap and may be correlated. In this paper, the agile aspects are operationalized by agile values, which focus more on individuals and interactions, working software, customer focus, and responding to changes but do not entirely ignore processes and planning elements.

Agile methods may not have significant upfront planning and control processes but have tacit and smaller planning and control elements such as iteration planning, daily standups, iteration reviews, release planning, and product road mapping (VersionOne, 2016). The plan-driven aspects focus on processes to manage customer expectations, contracts, controls on schedule, cost and quality, scope creep, and risk. However, a plan-driven approach does not preclude iterative development or responding to changes such as in a RUP-like method (Kruchten, 2000).

For assessing the relative strengths of the relationships of agile values and plan-driven aspects with project success, a questionnaire survey-based study was conducted. In addition to agile values, plan-driven aspects, and project success, the quantitative research reported in this paper considered all antecedents from the (Batra, 2017) qualitative research except organizational culture, which had some confounding issues with agile values during Q-sort analysis conducted before collecting the survey data. This section presents the literature used to develop the questionnaire, which is included in Appendix A. All figures in the paper are extracted from analyses conducted using SmartPLS version 3 (Ringle, Wende, & Becker, 2015). Figure 1 shows the proposed research model with the key indicators. After the initial data analysis, one indicator (scope creep) was dropped because of the reliability issue. The survey includes project success as the outcome variable, agile values, and plan-driven aspects as the mediating variables, and technological capability, shared understanding, top management commitment, and complexity as antecedent variables.

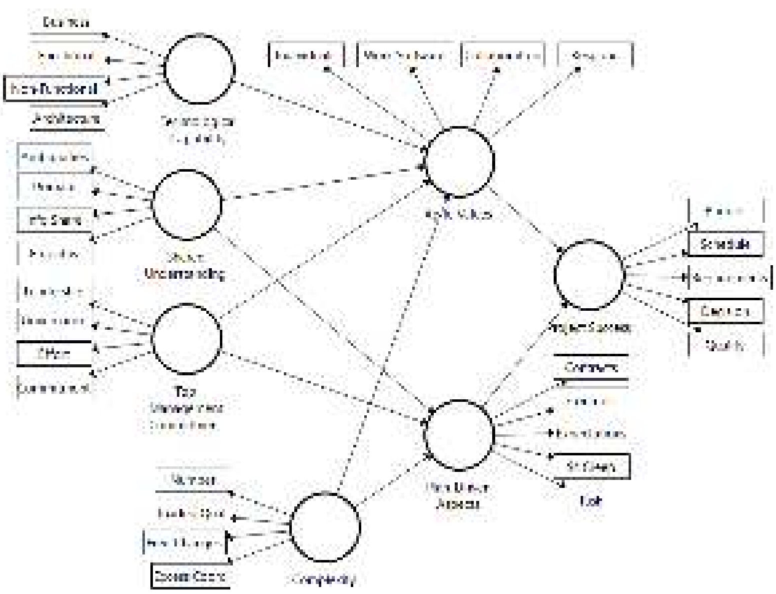

FIGURE 1: INITIAL STRUCTURAL AND MEASUREMENT MODEL FOR RESEARCH

2.1 Project Success

Project success refers to the measurement related to project implementation and efficiency, such as cost, duration, specification, and process efficiency while user satisfaction is an attitudinal measure that reflects meeting customer requirements on quality, effectiveness, and efficiency (Siau, Long, & Ling, 2010). Studies usually consider cost, schedule, quality, and user satisfaction as critical measures of project success (Heck & Zaidman, 2016; McLeod, Doolin, & MacDonell, 2012; Siau et al., 2010). In the context of DW/BIA, another measure of project success is decision-making; in fact, DW/BIA can be traced back to decision support systems in the 1980s. Sharda et al. (2016) and Davenport and Harris (2017) assert that managers, especially those at high managerial levels, are primarily decision makers, and are supported by DW/BIA systems, which enhance data management, analytical support, cognitive support, and knowledge management. Thus, the measures of project success included items on budget, schedule, quality, customer requirements, and decision making.

2.2 Agile Values

The common thread in agile development is the Agile Manifesto (Fowler & Highsmith, 2001) with its four values and the related twelve principles. The various agile principles, practices, and methods primarily stem from the Agile values (Boehm & Turner, 2004). These values are relevant even in scaled agile although not all principles are relevant (Dikert, Paasivaara, & Lassenius, 2016; Knaster & Leffingwell, 2017). Agile development has been shown to be especially useful when the response to changes and customer collaboration are essential (Boehm & Turner, 2004). A working DW/BIA allows early failure thus averting large-scale collapse, and interactions between the users and the developers facilitate value-driven development (Collier, 2012). The four values focus on individuals and interactions, working software, customer collaboration, and responding to changes and were used to develop the construct Agile Values. Adoption of agile values was hypothesized to have a positive effect on project success (hypothesis 1).

The agile values construct is used as a surrogate of agile aspects. For reducing measurement error, it is customary to have about four to six indicators for a construct. The Agile Manifesto lists four agile values, which is a reasonable number to measure agile aspects. For the sake of parsimony, the many agile principles and practices were not included. There was no guarantee that a majority of respondents would be using specific agile methods like Scrum or agile practices such as pair programming. It was possible that respondents could be using company-specific methods based on some agile values but not on specific agile practices.

2.3 Plan-Driven Aspects

A plan-driven approach has process management aspects that control schedule, costs, and quality (Humphrey, 1995; Kan, 2002). Project and process management constitute the processes of managing the achievement of the project objectives (Munns & Bjeirmi, 1996) and involve planning, control, and risk management (Boehm & Turner, 2004; Daneva et al., 2013) to manage customer expectations (Van Waardenburg & Van Vliet, 2013) and to fulfill contracts and promises (Batra, 2017). Requirements determination is a critical step in any software systems development (McLeod & MacDonell, 2011). Scope creep can result in project failure (Wallace & Keil, 2004). A contract enhances success especially in large or outsourced projects (Boehm & Turner, 2004; Sabherwal, 1999). Establishing control processes increases the likelihood of achieving success in software projects (Dingsoyr, Nerur, Balijepally, & Moe, 2012; Mao, Lee, & Deng, 2008; Misra, Kumar, & Kumar, 2009; Turk, France, & Rumpe, 2005). A risk management strategy augments the success of a project (Boehm, 1991; Wallace & Keil, 2004). Thus, scope creep, expectations management, contracts, controls, and risk management served as primary measures of plan-driven approach that could have a positive effect on project success (hypothesis 2).

2.4 Shared Understanding

Communication increases shared understanding (Conboy & Morgan, 2011) and although the two may be considered synonymous, communication can be viewed as the process that leads to a shared understanding. The reduced documentation in agile development increases the need for tacit knowledge, which necessitates a shared understanding using communication mechanisms such as standup meetings (Conboy & Morgan, 2011) and good quality social interactions (Coughlan, Lycett, & Macredie, 2003; Ryan & O'Connor, 2013). In agile software development projects, communication is a critical factor (Hummel, Rosenkranz, & Holten, 2013; Karhatsu, Ikonen, Kettunen, Fagerholm, & Abrahamsson, 2010; Mishra, Mishra, & Ostrovska, 2012) given that agile development is people-centric and needs collaboration for collective action (Batra, Xia, & Rathor, 2016; Cockburn & Highsmith, 2001; Nerur et al., 2005).

Communication is also essential for plan-driven aspects (Hwang & Ng, 2013; Papke-Shields, Beise, & Quan, 2010) which is a necessary component of the PMBOK guide (Rose, 2013). Documented knowledge is crucial to requirement specifications, drafting contracts, control mechanisms, and risk management. Domain knowledge, whether tacit or recorded, is a vital asset for setting up successful client-developer collaboration (Daneva et al., 2013) especially in the DW/BIA environment (Batra, 2017). In this study, a four-item scale based on removing ambiguities, information sharing, empathic dialogue, and domain knowledge was used to measure the construct shared understanding, which was hypothesized to improve the achievement of both agile values and plan-driven aspects (hypotheses 3 and 4).

2.5 Top Management Commitment

Top management commitment may be critical for management of DW/BIA projects (Batra, 2017; Davenport & Harris, 2017). The lack of management commitment is an obstacle to implementing software development (Huisman & Iivari, 2002). Top management commitment is essential for agile development (Senapathi & Srinivasan, 2012) and plan-driven approach (McLeod & MacDonell, 2011). Management apathy is a crucial challenge in the adoption of agile development (Vijayasarathy & Turk, 2008). Thus, effort and commitment from top management are essential. Leadership (Knaster & Leffingwell, 2017) and governance (Batra, 2017) are other critical top management commitment elements. In this study, a four-item scale, which addressed effort, commitment, leadership, and governance, was used to measure the construct top management commitment, which was hypothesized to positively affect the constructs agile values and plan-driven aspects (hypotheses 5 and 6).

2.6 Technological Capability

The Agile Manifesto recognizes the role of individuals as an essential value and references technical and design excellence as agile principles (Fowler & Highsmith, 2001). Team capability is a significant factor for agile development (Chow & Cao, 2008; Maxwell & Forselius, 2000). Agile development does not work unless there is a sufficient number of technologically capable personnel (Boehm & Turner, 2004) and the team has the delivery capability (Chow & Cao, 2008). Developer capabilities and experience are influential factors in gaining confidence, resolving experience-related risks, reducing implementation level code breakage, and faster resolution of design deficiency (Tan et al., 2009). Technological capability focuses not only on business and functional requirements, but also on non-functional requirements (Ramesh, Cao, & Baskerville, 2010). The team should be able to deliver business, functional, non-functional, and architecture requirements (Batra, 2017; Collier, 2012; Knaster & Leffingwell, 2017). These four indicators were used to develop a scale to measure the construct technological capability, which was hypothesized to have a positive effect on the construct agile values (hypothesis 7). The technological capability construct was not hypothesized to affect plan-driven aspects.

2.7 Complexity

The complexity of a project can stem from a number of factors such as large size (Moe, Aurum, & Dyba, 2012; Vlietland & van Vliet, 2015), integration, quality, coordination challenges and reconciling viewpoints of a wide variety of stakeholders (Batra, 2017), and the requirements and technological changes (Xia & Lee, 2003). There are two essential components of software complexity: a large number of parts that denote structural complexity and changing requirements that indicate dynamic complexity. Complexity should have a positive effect on plan-driven aspects because a large number of elements (i.e., structural complexity) require planning, coordination, and control. Complexity should also have a positive effect on agile values because agile development is specifically geared to respond to changes (i.e., dynamic complexity).

Complexity in DW/BIA can arise because of other factors such as inadequate quality of incoming data (Sharda et al., 2016) and the excess coordination effort among the stakeholders (Knaster & Leffingwell, 2017). In a larger project, the product owner needs to consider the viewpoints of a wide variety of stakeholders, many of whom may have conflicting views on the desirability of the software features and its functionality (Moe et al., 2012; Vlietland & van Vliet, 2015). In this study, a large number of pieces during integration, inadequate quality of incoming data, frequent changes in requirements, and excess coordination efforts formed a four-item scale to measure the construct complexity, which was hypothesized to positively affect agile values and plan-driven aspects (hypotheses 8 and 9).

2.8 Path Model and Hypotheses

The path model with structural relationships among the constructs is shown in Figure 1 shows while the complete item descriptions of the indicators are listed in Appendix A. The constructs agile values, and plan-driven aspects are the mediating variables between antecedents and project success. Shared understanding, technological capability, top management commitment, and complexity are the antecedent variables.

Based on the research instrument development and the path model, the following hypotheses were proposed:

H1) Agile values will have a positive effect on Project Success
H2) Plan-driven aspects will have a positive effect on Project Success
H3) Shared Understanding will have a positive effect on Agile values
H4) Shared Understanding will have a positive effect on Plan-driven aspects
H5) Top management commitment will have a positive effect on Agile values
H6) Top management commitment will have a positive effect on Plan-driven aspects
H7) Technological capability will have a positive effect on Agile values
H8) Complexity will have a positive effect on plan-driven aspects
H9) Complexity will have a positive effect on agile values

3. Methodology

The research model was evaluated based on the partial least squares for structural equation modeling (PLS-SEM) approach (Richter, Cepeda, Roldán, & Ringle, 2015) using SmartPLS version 3 software (Ringle et al., 2015). The PLS-SEM approach was selected because of the following reasons:

a) The purpose of the study was to investigate further the agile-planning framework proposed by (Batra, 2017). PLS-SEM is preferred when conducting exploratory research (Gefen, Straub, & Boudreau, 2000; Vinzi, Trinchera, & Amato, 2010).

b) The research objective was prediction rather than the confirmation of structural relationships (Hair, Ringle, & Sarstedt, 2011). At this point, there is inadequate theory regarding the research questions. It was essential to determine the explained variance (R-square) of the endogenous latent variables and the strength of the relationships (Hair, Sarstedt, Ringle, & Mena, 2012).

c) The sample size of 124 was consistent with the PLS recommendations (Hair, Hult, Ringle, & Sarstedt, 2016).

d) Multivariate normal data could not be guaranteed. PLS-SEM makes practically no assumptions about the underlying data (Cassel, Hackl, & Westlund, 1999).

The use of the alternative, covariance-based SEM (CB-SEM) approach, is appropriate for confirmatory research. CB-SEM requires a set of assumptions that include the multivariate normality of data and preferably a larger sample size (Diamantopoulos & Siguaw, 2000; Hair et al., 2011). Thus, the study did not employ CB-SEM.

PLS can be misapplied, and more attention should be paid to the assumptions of the PLS model (Rönkkö & Evermann, 2013; Rönkkö, Parkkila, & Ylitalo, 2012) although some of the alleged shortcomings have been refuted (Henseler et al., 2014; Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016). An alternative SEM technique based on artificial neural network (ANN) indicates results that are more comparable to PLS than to CB-SEM (Hsu, Chen, & Hsieh, 2006). The universal structure modeling (USM) approach, which is an SEM technique based on Bayesian neural network, reveals that the results are similar to both PLS and CB-SEM as long as linear relationships are assumed (Buckler & Hennig-Thurau, 2008). This section provides details on both the measurement model and the structural model analysis and includes the care taken to avoid inappropriate use of PLS.

Data were collected from 124 respondents, which is an adequate number based on power analysis (Hair et al., 2016) and is further explained in the data collection section later. The indicator data did not exhibit large skewness or kurtosis; most values fell in the recommended -1 to 1 range. This section summarizes the following checks for the quality and validity of data: Q-sort analysis, sample size, indicator reliability, and convergent validity (see Table 1) and discriminant validity (see Table 2), common method bias, and other minor validations. The analysis followed the guidelines prescribed in (Hair et al., 2016). Tables 3, 4, and 5 show the distribution of respondents by the industry types, the methodology, and the respondent roles. For collecting data, the researcher developed the questionnaire using the Qualtrics software and hired the Qualtrics Software Company, which provided data from 108 respondents. Additionally, the researcher posted the survey on his LinkedIn account and received 16 responses.

3.1 The Partial Least Square Approach

The overall purpose of PLS-SEM is to minimize the amount of unexplained variance to predict the dependent variable. PLS-SEM has two steps: 1) to validate the measurement model, and 2) to assess the strength of relationships in the structural model (Hair et al., 2016). The relationships between a construct and its measures constitute the measurement (or outer) model and the relationship among the latent constructs is called the structural (or inner) model (Hair et al., 2011). The measurement model depicts how indicators (or items) accurately measure latent constructs such as perceived usefulness, attitude, self-efficacy, shared understanding, or commitment. A latent construct is more challenging to measure because it is abstract and not directly observable as a concrete measure such as income or number of cars sold. Thus, a latent construct is measured with a set of indicators that serve as proxy variables (Hair et al., 2016). The measurement model is validated first by performing various validity checks such as reliability, and convergent and discriminant validity.

After the measurement model has been validated, the strengths of relationships in the structural model are estimated. The structural model is proposed in advance by examining the theoretical sources, literature, or a qualitative study anchored in grounded theory. The PLS-SEM algorithm uses standardized data and calculates standardized coefficients between -1 and 1. A higher absolute value of a coefficient represents a stronger relationship. A bootstrapping procedure is used to estimate standard errors and assess if a given coefficient is significant at a 5% level. The coefficient is expressed by the notation a (b); for example, 0.305 (0.002) means that the standardized coefficient is 0.305 and the significance level is 0.002 or 0.2%.

A detailed treatise of PLS-SEM is available in the book by Hair et al. (2016). Based on PLS-SEM, SmartPLS 3 (Ringle et al., 2015) is a statistical software frequently used for determining the strength of simultaneous relationships among variables/constructs (and especially among latent constructs). As may be noted in Figure 1, the convention is to display the PLS-SEM model by using circles or ovals for constructs and small rectangles for the indicators (or items). A construct is usually measured by at least four indicators to reduce the measurement error. A given indicator can be measured on a five- or seven-point Likert scale.

When using SmartPLS, the sample size does not have to be large, but statistical power tables show that samples sizes below 100 are rare unless the R-square is at least moderate. Based on (Cohen, 1992), Hair et al. (2016) provide a convenient statistical power table, which allows a researcher estimate the sample size based on 1) minimum R-square values of 0.10, 0.25, 0.50, and 0.75 (lower value requires a higher sample size); 2) significance level of 1%, 5%, and 10% (i.e., type 1 error) 3) statistical power of 80% (i.e., a type 2 error of 20%). For a significance level of 5%, six independent variables, and minimum R-square of 0.25, the recommended sample size is 48; if the minimum R-square is 0.10, then the recommended sample size is 130. This study had a sample size of 124 which corresponds to the conservative level of R-square.

3.2 Data Collection

The previous section explained how the constructs and the indicators were derived. The indicator to construct mapping was tested by conducting a Q-sort analysis that involved three Information Systems Master's students who had experience in software development. The mapping revealed that one student was confused in differentiating the cards for agile values and organizational culture because the indicators of the latter construct were related to an adaptable culture. The concern was valid, and consequently, the construct organizational culture was dropped from the study. The questionnaire was then compiled using the Qualtrics software.

The Qualtrics Software Company was hired to conduct the data collection. Qualtrics assigned a consultant to manage the data collection, to add quality checks in the questionnaire, make recommendations, and screen out bad data, which was primarily because of respondent using the same choice for all of the indicators or filling out the questionnaire too quickly. Out of the total valid 124 responses, Qualtrics provided 108 responses while the researcher personally collected 16 responses thorough LinkedIn postings.

3.3 Reliability

Composite reliability was used to measure internal consistency across items within a construct. The values ranged from 0.85 to 0.93 and were significant at p-value = 0.000 (Table 1). The recommended range for composite reliability is 0.8-0.9 (Hair et al., 2016). Cronbach's alpha is another measure of reliability; these values ranged from 0.769 to 0.904.

3.4 Convergent Validity

Convergent validity is the extent to which a measure correlates positively with other measures of the same construct (Hair et al., 2016). In assessing convergent validity, indicators of a construct are treated as alternative approaches for measuring the same construct by calculating loadings. Indicator reliability was estimated by calculating the outer loadings, which ranged from 0.76 to 0.90 but mostly from 0.8 to 0.9. An outer loading less than 0.7 is usually not acceptable (Hair et al., 2016). The data revealed two values lower than 0.7. An indicator, scope creep, of the plan-driven aspects construct had an outer loading of 0.65 and was dropped. The data was analyzed again. Another indicator - commitments such as contracts or promises - of the plan-driven aspects indicator had a borderline outer loading of 0.683 and was close enough to 0.7 to justify inclusion; furthermore, it was deemed as theoretically necessary. All outer loadings were significant at p-value=0.000. Thus, the indicator reliability of the scales was established.

TABLE 1: RELIABILITY AND CONVERGENT VALIDITY

Construct	Composite Reliability	Average Variance Extracted (AVE) for Convergent Reliability
Agile Values	0.887	0.664
Complexity	0.891	0.673
Plan-Driven Aspects	0.852	0.592
Project Success	0.928	0.722
Shared Understanding	0.921	0.746
Technological Capability	0.914	0.727
Top Management Commitment	0.918	0.736

A standard measure to establish convergent validity on the construct level is the average variance extracted (AVE), which is the mean value of the squared loadings of the indicators associated with the construct (Hair et al., 2016). The recommended minimum AVE value is 0.5, which indicates, that on average, the construct explains more than 50% of the variance of its indicators. The analysis revealed that the values were between 0.592 and 0.746. Thus, acceptable AVE values demonstrate the convergent validity of the constructs.

3.5 Discriminant Validity

Discriminant validity is the extent to which a construct is truly distinct from other constructs by empirical standards (Hair et al., 2016). One approach to establishing discriminant validity is by using the Fornell-Larcker criterion, which compares the square root of the AVE values with the latent variable correlations. However, the Fornell-Larcker has some limitations (Voorhees, Brady, Calantone, & Ramirez, 2016) and a new approach, heterotrait-monotrait ratio (HTMT), is considered as an improvement (Henseler, Ringle, & Sarstedt, 2015). An HTMT value above 0.9 demonstrates a lack of discriminant validity. As indicated in Table 2, all HTMT ratios are below 0.9, which establishes the discriminant validity among the constructs.

TABLE 2: DISCRIMINANT VALIDITY USING HTMT RATIO

	Agile Values	Complexity	Plan-Driven Aspects	Project Success	Shared Understanding	Technological Capability	Top Management Commitment
Agile values							
Complexity	0.349						
Plan-driven aspects	0.692	0.356					
Project Success	0.683	0.179	0.558				
Shared Understanding	0.685	0.456	0.681	0.419			
Tech. Capability	0.654	0.401	0.657	0.485	0.776		
Top Manage Commit	0.565	0.277	0.818	0.383	0.609	0.692	

3.6 Common Method Bias

The un-rotated factor analysis using all latent constructs was performed to check if a single factor emerged that explained the majority of the variance in the model, which would have indicated the common method bias (Lowry & Gaskin, 2014). The analysis showed multiple factors; the highest variance explained by a factor was 37.5%. The results suggested that the data was not confounded by the common method bias. Note that the respondents did not belong to one or a few organizations but from a large number of organizations. This diversity is reflected in the analysis of respondent distribution.

3.7 Respondent Distribution

The sequence of questions in the survey ensured that the respondents fulfilled some initial criteria for selection. The respondents were asked if they would complete the survey to the best of their knowledge; if they selected any other option, they were excluded and taken to the end of the survey. Furthermore, they were asked about the job type. If the job type was not one of data warehousing, analytics, or business intelligence development, they were also excluded. The distribution of respondents by job type was 42% for Analytics and 29% each for Data Warehousing and Business Intelligence. The median project size regarding team members was 15. Three other questions were asked to gain a better perspective of the respondent distribution: industry, method, and role. Table 3 summarizes the respondent distribution by industry.

TABLE 3: RESPONDENT DISTRIBUTION BY INDUSTRY

Industry Name	Respondent Industry (Percent)
Education/Research	15.8
Manufacturing	14.9
Healthcare	13.2
Marketing	10.5
Financial Services	7.9
Telecommunication	6.1
Media/Entertainment	4.4
Transportation	4.4
Other	22.8

Table 4 indicates the respondent distribution by the methodology used. Respondents seem to be mainly using company-owned methods. The three agile methods - Scrum, Kanban, and Scaled Agile - put together constituted 17.7% and would rank second if these differentiated choices are combined. It appears that named agile methods are not widely being used in DW/BIA development but this does not mean that agile values are not considered necessary in company-owned or hybrid methods. Many organizations do not use named methods and instead rely on ad-hoc approaches based on the useful aspects of the methodologies (Avison & Fitzgerald, 2003).

TABLE 4: RESPONDENT DISTRIBUTION BY THE METHODOLOGY

Method Name	Method (Percent)
Company Owned	40.4
Hybrid	14.9
Waterfall	9.6
Rational Unified Process	8.8
Scaled Agile	8.8
Kanban	4.4
Scrum	3.5
Other	9.6

Table 5 shows the respondent distribution by role. The two dominant roles were DW/BIA analyst and manager. This distribution was not surprising as the analysts and managers are frequently involved in DW/BIA development.

TABLE 5: RESPONDENT DISTRIBUTION BY ROLE

Role Name	Role (Percent)
DW/BIA Analyst	27.2
Manager	24.6
Administrator	9.6
Consultant	9.6
Developer	7.0
Systems Analyst	7.0
Other	14.9

4. Results

4.1 Initial Structural Model Results

After the measurement model was validated, the structural model was tested using the software SmartPLS version 3 (Ringle et al., 2015). The path coefficients and significance (p-values) were obtained using 5000 runs of PLS algorithm and bootstrapping. The path coefficients, which represent the strength of the hypothesized relationships among the constructs, have standardized values approximately between -1 and +1 (Hair et al., 2016). Whether or not a coefficient is significant ultimately depends on its standard error. Sample means are necessary for obtaining the standard error. The software employs an approach called bootstrapping, which draws random samples from the data to get the sample means and determine the standard error of the magnitude of each relationship, which may be between two constructs or between a construct and its indicator.

A significance level of 0.05, which indicates that there is a 5% probability that a coefficient or a loading is significant by chance, was used in this study. Figure 2 shows the initial results of the analysis. Note that the indicator, scope creep, has been removed from the figure. The p-values for the indicators were all significant at p=0.000 and are not shown. Except in the case of complexity, all other path coefficients were consistent with the direction of the hypotheses, with four out of eight being significant at the 5% level. The direct effect of each of the four antecedents – technological capability, shared understanding, top management commitment, and complexity – to project success was negligible; thus, agile values and plan-driven aspects mediated the relationships from the antecedents.

The coefficient of determination (R-square) was moderate for all three endogenous constructs: 0.41 for Agile values, 0.51 for Plan-driven aspects, and 0.39 for Project success. Given the R-square values, the sample size of 124, which was estimated at R-square of about 0.10, is adequate. However, the results had many marginally significant coefficients, and it seemed that given the relatively exploratory nature of the study, further investigation was necessary. The second stage of analysis is detailed in a later section.

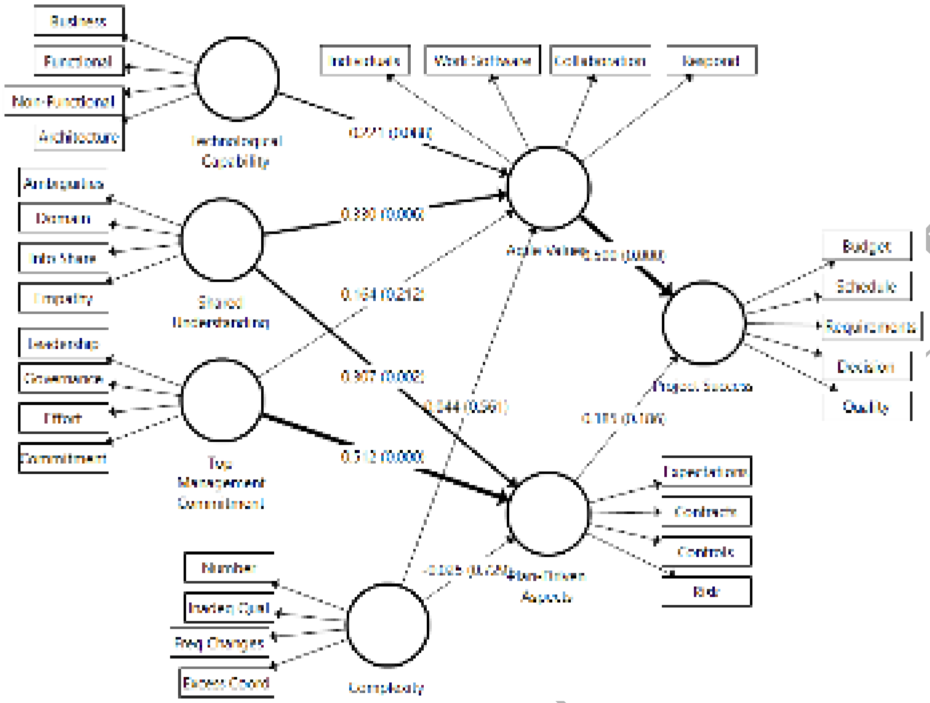

FIGURE 2: COEFFICIENT LEVELS AND SIGNIFICANCE LEVELS IN THE PATH MODEL

4.2 Initial Hypotheses Significance

Hypothesis H1, agile values will have a positive effect on project success, was strongly supported. The path coefficient 0.50 was strong and significant at a level of 0.000. In the DW/BIA domain, we can unequivocally claim that agile values are the key to project success. Hypothesis H2, plan-driven aspects will have a positive effect on project success, was not supported. The path coefficient 0.185 was in line with the direction of the hypothesis, but it was too low for statistical significance (p=0.186). Hypotheses H3, shared understanding will have a positive effect on agile values, and H4, shared understanding will have a positive effect on plan-driven aspects, were both supported (p=0.006 and p=0.002) and had similar coefficients (0.33 and 0.31). Hypothesis H5, top management commitment will have a positive effect on agile values, was not supported. Hypothesis H6, top management commitment will have a positive effect on plan-driven aspects, was strongly supported with a path coefficient of 0.51 and a p-value of 0.000. With a path coefficient of 0.22 and a p-value of 0.088, Hypothesis H7, the technological capability will have a positive effect on agile values, was found not significant at p=0.05. Hypotheses H8 and H9, which addressed complexity, had weak effects and were not significant.

Overall, the initial results did not appear to be clear because it seemed that plan-driven aspects, top management commitment, and technological capability contributed to the model but were not statistically significant. Furthermore, the results were inconsistent as compared with the qualitative findings of the (Batra, 2017) study. It seemed like a routine PLS-SEM analysis did not reveal the deeper relationships because of potential observed or unobserved heterogeneity of data. Thus, further analysis was conducted by first examining observed heterogeneity on job type and moderation effects of complexity on agile values and plan-driven aspects. None of these results were significant. Finally, the data were examined for unobserved heterogeneity. The results indicated that there were two underlying segments with distinctly different characteristics.

4.3 Hypotheses Significance after considering Unobserved Heterogeneity

Unobserved heterogeneity, which pertains to the existence of more than one subpopulation that is not distinguished by a previously identified variable, can threaten different types of validity (Becker, Rai, Ringle, & Völckner, 2013). To discover unobserved heterogeneity in both structural and measurement models, Becker et al. (2013) propose a new method – prediction-oriented segmentation (PLS-POS) – to overcome the limitations of other distance measure-based methods. Standard clustering methods such as k-means clustering focus only on indicator data when forming groups of data but they cannot account for latent variables and their structural model relationships (Hair et al., 2016; Sarstedt & Ringle, 2010). The PLS-POS method can detect heterogeneity in the measurement and the structural model and does not require that data be normally distributed. The study preferred PLS-POS because an alternative approach, FIMIX-PLS (Sarstedt, Becker, Ringle, & Schwaiger, 2011), requires that the endogenous variables have a multivariate normal distribution, which is inconsistent with the distribution-free assumption of PLS.

The initial results had revealed many coefficients that had low to moderate strength but were not significant. It was reasonable to assume that the results could stem from two conflicting subpopulations. The initial results established the significance of agile values. Based on Boehm and Turner (2004) recommendations, there was theoretical support for a balanced development method, which would give similar weight to both agile values and plan-driven aspects. Thus, PLS-POS was run with two segments because it was likely that the balanced segment was being mitagated by the agile-heavy segment. The measurement model was sound except that one indicator – risk – had to be dropped. For avoiding repetition, the measurement model is not shown in subsequent analysis; note that the [+] in each circle indicates that the construct has indicators. Figure 3 shows the result of the agile-plan balanced segment, which accounted for 85 of the 124 cases. Figure 4 shows the result of the agile-heavy segment, which accounted for 39 of the 124 cases. The analysis revealed that the two segments represented two subpopulations that seemed to be mitigating each other’s effects.

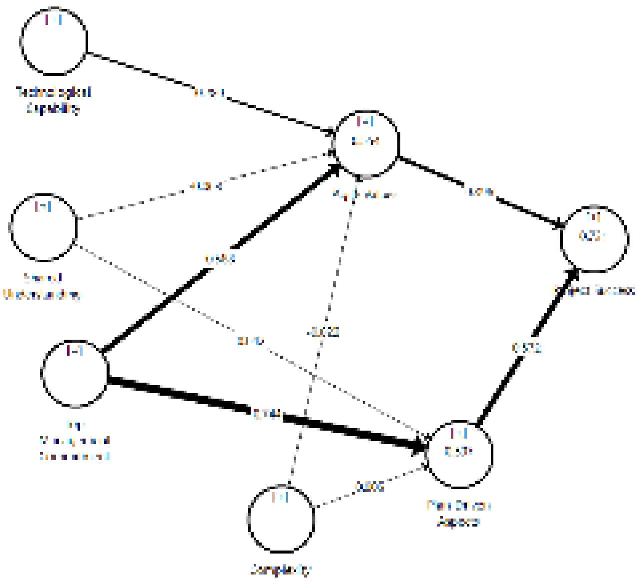

FIGURE 3: THE AGILE-PLAN BALANCED SEGMENT

Table 6 indicates that the agile-plan balanced segment shows relatively strong coefficients for both the relationships between agile values and project success, and plan-driven aspects and project success. The R-square is very strong. Out of the nine hypotheses, five are supported for the agile-plan balanced segment. Both agile values and plan-driven aspects have a complementary and robust effect on success. Top management commitment has strong impact on agile values, and plan-driven aspects whereas shared understanding and complexity have negligible effects. The technological capability construct has a moderate and significant effect on agile values. The direct effect of each of the four antecedents – technological capability, shared understanding, top management commitment, and complexity – to project success was negligible, and these results are not included.

TABLE 6: HYPOTHESES SUPPORT OF THE BALANCED SEGMENT

Path	Path Coefficient	p-Value	Hypothesis
Agile Values -> Project Success	0.396	0.000	Supported
Plan-Driven Aspects -> Project Success	0.572	0.000	Supported
Complexity -> Agile Values	-0.022	0.763	Not Supported
Complexity -> Plan-Driven Aspects	0.005	0.950	Not Supported
Shared Understanding -> Agile Values	-0.063	0.589	Not Supported
Shared Understanding -> Plan-Driven Aspects	0.042	0.641	Not Supported
Technological Capability -> Agile Values	0.28	0.030	Supported
Top Management Commitment -> Agile Values	0.588	0.000	Supported
Top Management Commitment -> Plan-Driven Aspects	0.744	0.000	Supported

As Figure 4 and Table 7 indicate, the agile-heavy segment showed a strong positive coefficient from agile values to project success and a strong negative coefficient from plan-driven aspects to project success. Out of the nine hypotheses, five are supported, and another two are significant but have a negative coefficient. The shared understanding construct is strongly related to both agile values and plan-driven aspects, whereas top management commitment is negatively and strongly related to agile values. Complexity is significantly associated with agile values but not to plan-driven aspects. The coefficient between technological capability and agile values is borderline, but because the p-value is 0.055, it is rounded off and deemed as significant. The direct effect of each of the four antecedents – technological capability, shared understanding, top management commitment, and complexity – to project success was negligible, and these results are not included.

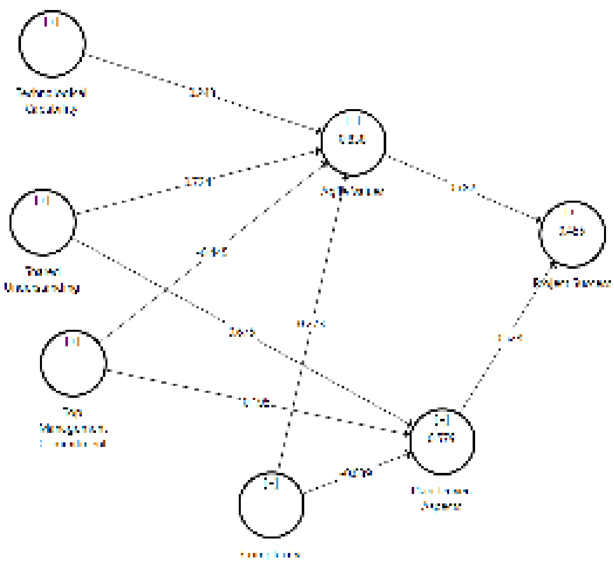

FIGURE 4: THE AGILE-HEAVY SEGMENT

TABLE 7: HYPOTHESES SUPPORT OF THE AGILE-HEAVY SEGMENT

Path	**Path Coefficients**	**p-Value**	**Hypothesis**
Agile Values -> Project Success	0.782	0.000	Supported
Plan-Driven Aspects -> Project Success	-0.528	0.005	Significant (reverse sign)
Complexity -> Agile Values	0.273	0.000	Supported
Complexity -> Plan-Driven Aspects	-0.039	0.819	Not Supported
Shared Understanding -> Agile Values	0.724	0.000	Supported

Shared Understanding -> Plan-Driven Aspects	0.675	0.000	Supported
Technological Capability -> Agile Values	0.243	0.055	Supported
Top Management Commitment -> Agile Values	-0.445	0.000	Significant (reverse sign)
Top Management Commitment -> Plan-Driven Aspects	0.195	0.279	Not Supported

By using a multigroup test, the path coefficient differences between the balanced and agile-heavy segments were examined for significant differences. Except for the paths complexity -> plan-driven aspects and technological capability -> agile values, the remaining seven paths were found to significantly different ($p < 0.05$). The significance of the differences supports the information provided by the diagrams that the two groups are indeed distinct and belong to different subpopulations.

5. Discussion, Limitations, and Future Research

5.1 Discussion

The use of the agile approach is mainly supported by practitioner literature such as (Collier, 2012; Hughes, 2012), which have recommended explicitly that DW/BIA development should consider the agile approach. Boehm and Turner (2004) proposed the notion of harmonizing agility and discipline in the same development approach. This study hypothesizes agile values and plan-driven aspects as constructs that simultaneously affect project success for DW/BIA development. The study attempted to answer two research questions: 1) Which factor, agile values or plan-driven aspects, contributes more toward the success of DW/BIA? 2) What are the significant antecedents of agile and plan-driven aspects?

Overall, the results indicate that agile aspects are more important than plan-driven aspects. Furthermore, the results suggest that the answers need to be examined in light of the two subpopulations, agile-plan balanced and agile-heavy, which were identified after the PLS-POS analysis. The respondents in the agile-plan balanced segment accounted for 68.5% and those in the agile-heavy segment accounted for 31.5% of the sample. The balanced configuration is consistent with the maxim proposed in (Boehm & Turner, 2004) and recent methods such as SAFe (Knaster & Leffingwell, 2017), LeSS (Larman & Vodde, 2016), and other large-scale agile methodologies (Ambler, 2009) that recommend that agile methods and plan-driven aspects need to be aligned.

Conversely, the agile-heavy segment, which constitutes a smaller proportion of data, strongly eschews planning and strictly follows agile values. Overall, the agile values construct is more critical because it indicates significance in both segments and the essential contribution of this research is the shift in the debate from agile versus structured to agile-heavy versus agile-plan balance. In the analytics domain, the plan-only approach would be rare if not altogether absent.

The significance of antecedents for agile values and plan-driven aspects in DW/BIA is mainly dependent on the segment. In the case of the balanced segment, top management commitment is the dominant antecedent. This finding makes sense because the senior management generally prioritizes planning, contracts, and controls; however, in this scenario, the senior management concomitantly emphasizes agile values, which tacitly represent an innovative and iterative manner of working. The top management is not only using the DW/BIA product for making decisions, but it seems to be involved in the development process. Thus, the top management in the DW/BIA domain does not depict the traditional command and control manner of support. Practitioners may especially note that top management has an enormous role in the balanced scenario. Top management leadership and commitment is especially important for issues such as acceptance and implementation of analytics recommendations (Davenport & Harris, 2017).

In the agile-heavy segment, top management commitment is not related to plan-driven aspects and is negatively associated with agile values. It is understandable that top management commitment is not a principle in the Agile Manifesto and may not be an essential factor for a purely agile development; nevertheless, the negative relationship suggests a different organizational culture that is wary of top management. This negative relationship needs further investigation in the future research using a corporate culture lens.

In the agile-heavy segment, shared understanding has a substantial effect on agile values. The strong emphasis on individuals and interactions, self-organization, customer feedback, face-to-face meetings, and tacit knowledge underscore the need for communication. Shared understanding is also considered important for plan-driven aspects, but this effect does not translate to project success because, for the agile-heavy segment, the construct plan-driven aspects has a negative relationship with project success. Shared understanding is not found to be an important antecedent for the balanced segment. One can conjecture that shared understanding is a proxy for self-organization, which is absent in the balanced segment because the top management may be an essential driver. Thus, we have two contrasting styles of development. The agile-heavy approach relies solely on agile values and employs shared understanding as a guiding vehicle while eschewing both planning and top management involvement. Conversely, the agile-plan balanced approach harmonizes agile values and plan and is driven by strong top management involvement and direction and less need for instituting processes that facilitate shared understanding.

One would think as complexity increases, the need for plan-driven aspects increases; however, complexity has almost a zero effect on plan-driven aspects in both balanced and agile-heavy segments. In the agile-heavy segment, complexity does have a moderate, significant effect on agile values. It is possible that the volatility of requirements may be the primary factor that causes complexity. Overall, complexity seems to be manageable. The availability of a plethora of technologies on scalable platforms especially on the cloud may be mitigating the effect of complexity. Furthermore, machine learning is augmenting human judgment to cope with complexity. As an example, Davenport and Harris (2017) describe a case on propensity modeling in which a company went from handling 150 models to 5000 models in three years through sustained improvement.

The effect of technological capability and agile values was marginal but significant in both segments. Perhaps, technological capability seems to be keeping pace and harnessing the increased complexity. A review of the evolving DW/BIA market indicates that there is undoubtedly some complexity in product development. Davenport and Harris (2017) list four stages of analytics, plausibly each more iterative and less structured than the other. Analytics 1.0 lasted until about 2005 and was characterized by reporting and decision support. In a short period, three more stages have emerged. Analytics 2.0 brought in the era of big data and technologies like Hadoop and NoSQL. Analytics 3.0 combined the previous two phases and employed proprietary and open-source technologies to focus on products and services that brought increased value to customers such as in job assistance feature offered by LinkedIn. Analytics 4.0, the current stage, is characterized by machine learning, which augments human knowledge with autonomous model evaluation and learning by machines to provide products and services to customers. Each generation subsumes the previous and then enhances some features. The move from structured to a more iterative approach in DW/BIA development may be explained by the rapid climb in analytics versions through business needs and technological advances that require and provide timely feedback. For example, the extension and refinement of models for better personalization necessitate an iterative and incremental delivery. The confirmation or refutation of the likely move from planning-oriented to balanced and agile-heavy modes is an area for future research.

The results from this study strongly support the use of agile values, which were used in the questionnaire to measure the underlying agile philosophy, but do not provide a clear indication of the use of specific agile methods. As shown in Table 4, the methodology choice is not dominated by specific agile methods. The three agile methods – Scrum, Kanban, and Scaled Agile – account for 17.7% whereas company specific methods constitute 40.4% and hybrid methods represent 14.9% (which may include agile hybrid methods). DW/BIA development is somewhat different from the conventional software development, which is dominated by coding. As may be noted from Table 5, developers constituted only 7% in the study whereas analysts constituted 27.2% of the respondents.

It is possible that at least in the DW/BIA arena, it is not about a specific methodology such as Scrum or Kanban, but more about the values underlying the methodology (Dikert et al., 2016). There may be an appeal to values such as customer collaboration, incremental development, responding to changes, and concern for individuals and interactions in a domain characterized by rapid changes and a quest for business value. It might just be that we have reached an era when such values are ingrained. Instead of analysts, managers, developers, and other participants using named agile methods, they have assimilated agile values into their development culture, whether agile-heavy or balanced. Perhaps, the methods are customized according to the context, but there is an underlying commitment to agile values and practices (Hoda, Kruchten, Noble, & Marshall, 2010; Murphy et al., 2013). The agile analytics book by Collier (2012) provides a list of agile methodologies but does not discuss any methodology in detail; instead, it lists "tenets of agility." The comprehensive textbook by Sharda et al. (2016) does not address the use of named agile or the structured development methods.

Contemporary development methods in DW/BIA seem to be compatible with agile values such as iterative development, responding to changes, focusing on the customer, and emphasizing individuals and interactions. Siegel (2016) discusses many case studies on predictive analytics involving iterative improvement, from Netflix's 2008 competition to improve the prediction of customer ratings to the IBM Watson's duel with the best contestants on the Jeopardy show. Davenport and Harris (2017) provide a similar sense of achieving incremental gains through model enhancement and offer some illustrative cases that include DW/BIA development at data-driven companies such as UPS, Google, Amazon, Progressive, LinkedIn, and Caesars. In each case study, there is a sense of iterative development that is explicitly or implicitly focused on customer value through punctuated or continual improvement. In a qualitative study (Batra, 2017), only one of the five companies was using a named agile method, but all respondents agreed that the agile way of working is compatible with DW/BIA. Contemporary DW/BIA development is not only consistent with but is dominated by agile values. However, the mismatch between the emphasis on agile values and the use of named agile methods deserves more investigation and is a topic for future research.

5.2 Limitations and Future Research

The study has some limitations, which point to future research avenues. The segmentation of the data into agile-plan balanced and agile-heavy is a strength of the research, but the effects of some antecedents are not entirely clear. For example, it is not clear why the impact of top management commitment on agile values is viewed quite negatively in the agile-heavy group, or why shared understanding has no role in the agile-plan balanced segment. The role of complexity is also somewhat unclear given that it has no part in the balanced segment but has some effect in the agile-heavy segment. These are issues for future research. The role of the current research should be viewed as exploratory especially because the analysis had to use PLS-POS to uncover heterogeneous subpopulations. In a prospective study, this distinction could be made deliberately so that a multigroup analysis can be conducted directly rather than in a post-hoc manner.

Overall, agile values are preferred over plan-driven aspects, but the definitive reasoning behind this finding is lacking. There is a possibility that respondents may have an attitude of favoring anything agile in the survey since it is widely perceived as contemporary and successful. In the study, there was no direct indication that the choices were part of the Agile Manifesto although some respondents might have recognized the items anyway. Because the study did not delve into more detail on the method used, the nature of company-specific and hybrid methods is unknown. Future research can try to understand the disconnection between the preference for agile values and the lack of adoption of specific, named agile methods. Furthermore, the nature of company-specific methods needs to be understood.

The research employed the PLS-SEM approach for data analysis but did not consider some alternative methods because of the lack of widely available software. A robustness testing of an artificial neural network (ANN)-based SEM technique indicates that it shows similar results as compared with both covariance (LISREL and EQS) and component-based (PLS) SEM techniques (Hsu et al., 2006); additionally, ANN-SEM can address non-linear relations between variables. A Bayesian neural network approach (Buckler & Hennig-Thurau, 2008) can be even more effective because it enables researchers to identify hidden structures within their models, unstipulated model paths, and nonlinear relations among model variables. Bayesian networks are popular for evidence-based decision-making in software engineering (Dejaeger, Verbraken, & Baesens, 2013; Misirli & Bener, 2014). However, Bayesian networks do not understand causes and effects because by design, in a Bayesian network, information flows in both directions (Pearl & Mackenzie, 2018). Commercial availability of advanced approaches that can handle both the measurement and structural models of latent variables can improve the theoretical framework and validation of DW/BIA development. Finally, the results need to be validated by using qualitative methods such as case studies in the real world environment.

6. Conclusion

The study investigated two research questions. 1) Which factor, agile values or plan-driven aspects, contributes more toward the success of DW/BIA? 2) What are the significant antecedents of agile values and plan-driven aspects? The results are not entirely straightforward because data analysis revealed two segments – agile-plan balanced and agile-heavy – with significantly distinct characteristics. Agile values are prominent in both approaches. In a way, the segments represent two separate organizational cultures as much as different development approaches. The balanced approach has top management commitment whereas the agile-heavy approach has shared understanding as the vital antecedent. The technological capability has a limited but significant effect on agile values. Complexity affects agile values but only in the agile-heavy case.

The reader should take special note on the process of emergence of agile-plan balanced and agile-heavy segments from unobserved heterogeneity. In future research, it may be apt to include a question that explicitly identifies the style of the development method. By using a separate measure of balanced versus agile-heavy style, the research can conduct a multigroup analysis based on observed heterogeneity, which would confirm the existence of the two groups. Furthermore, the research could consider questions on the organizational culture.

The current quantitative study was built upon a qualitative study. Some of the issues that have arisen from this study can be addressed by conducting multiple case studies whereas other matters may be better answered by further quantitative studies. However, the current study does provide a definite finding of the shifting of debate from agile versus structured to agile-plan balanced versus agile-heavy approaches.

References

Abbasi, A., Sarker, S., & Chiang, R. (2016). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems, 17*(2), i-xxxii.

Ambler, S. W. (2009). The agile scaling model (ASM): adapting agile methods for complex environments. *Environments*.

Ariyachandra, T., & Watson, H. (2010). Key organizational factors in data warehouse architecture selection. *Decision Support Systems, 49*(2), 200-212. doi:10.1016/j.dss.2010.02.006

Avison, D. E., & Fitzgerald, G. (2003). Where now for development methodologies? *Communications of the ACM, 46*(1), 78-82.

Batra, D. (2017). Adapting Agile Practices for Data Warehousing, Business Intelligence, and Analytics. *Journal of Database Management, 28*(4), 1-23.

Batra, D., VanderMeer, D., & Dutta, K. (2011). Extending agile principles to larger, dynamic software projects: a theoretical assessment. *Journal of Database Management (JDM), 22*(4), 73-92.

Batra, D., Xia, W., & Rathor, S. (2016). Agility Facilitators for Contemporary Software Development. *Journal of Database Management (JDM), 27*(1), 1-28.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*(3), 665-694.

Boehm, B. W. (1991). Software risk management: principles and practices. *IEEE Software, 8*(1), 32-41.

Boehm, B. W., & Turner, R. (2004). *Balancing agility and discipline: A guide for the perplexed*. Boston, MA: Addison-Wesley.

Buckler, F., & Hennig-Thurau, T. (2008). Identifying hidden structures in marketing's structural models through universal structure modeling. *Marketing JRM, 30*(2), 47-66.

Cassel, C., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of applied statistics, 26*(4), 435-446.

Chen, H., Chiang, R., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly, 36*(4), 1165-1188.

Chow, T., & Cao, D.-B. (2008). A survey study of critical success factors in agile software projects. *Journal of Systems and Software, 81*(6), 961-971.

Cockburn, A., & Highsmith, J. (2001). Agile software development, the people factor. *Computer, 34*(11), 131-133.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Collier, K. (2012). *Agile analytics: A value-driven approach to business intelligence and data warehousing*. Upper Saddle River, NJ: Addison-Wesley.

Conboy, K., & Morgan, L. (2011). Beyond the customer: Opening the agile systems development process. *Information and Software Technology, 53*(5), 535-542.

Coughlan, J., Lycett, M., & Macredie, R. D. (2003). Communication issues in requirements elicitation: a content analysis of stakeholder experiences. *Information and Software Technology, 45*(8), 525-537.

Daneva, M., van der Veen, E., Amrit, C., Ghaisas, S., Sikkel, K., Kumar, R., . . . Wieringa, R. (2013). Agile requirements prioritization in large-scale outsourced system projects: An empirical study. *Journal of Systems and Software, 86*(5), 1333-1353.

Davenport, T., & Harris, J. (2017). *Competing on Analytics: The New Science of Winning*: Harvard Business Press.

Dejaeger, K., Verbraken, T., & Baesens, B. (2013). Toward comprehensible software fault prediction models using bayesian network classifiers. *IEEE Transactions on Software Engineering, 39*(2), 237-257.

Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL: A guide for the uninitiated*. London, UK: SAGE.

Dikert, K., Paasivaara, M., & Lassenius, C. (2016). Challenges and success factors for large-scale agile transformations: A systematic literature review. *Journal of Systems and Software, 119*, 87-108.

Dingsoyr, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software, 85*(6), 1213-1221. doi:DOI 10.1016/j.jss.2012.02.033

Dyba, T., & Dingsoyr, T. (2008). Empirical studies of agile software development: A systematic review. *Information and Software Technology, 50*(9-10), 833-859.

Fowler, M., & Highsmith, J. A. (2001). The agile manifesto. *Software Development, 9*(8), 28-35.

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems, 4*(1), 7.

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage Publications.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice, 19*(2), 139-152.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414-433.

Heck, P., & Zaidman, A. (2016). A systematic literature review on quality criteria for agile requirements specifications. *Software Quality Journal*, 1-34.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., . . . Calantone, R. J. (2014). Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organizational Research Methods, 17*(2), 182-209.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135.

Hoda, R., Kruchten, P., Noble, J., & Marshall, S. (2010). Agility in context. *ACM Sigplan Notices, 45*(10), 74-88.

Hsu, S.-H., Chen, W.-h., & Hsieh, M.-j. (2006). Robustness testing of PLS, LISREL, EQS and ANN-based SEM for measuring customer satisfaction. *Total Quality Management & Business Excellence, 17*(3), 355-372.

Hughes, R. (2008). *Agile Data Warehousing: Delivering world-class business intelligence systems using Scrum and XP*. New York: IUniverse.

Hughes, R. (2012). *Agile data warehousing project management: business intelligence systems using Scrum*. Waltham, MA: Morgan Kaufmann.

Huisman, M., & Iivari, J. (2002). The individual deployment of systems development methodologies. In *CAISE* (pp. 134-150). Berlin: Springer-Verlag.

Hummel, M., Rosenkranz, C., & Holten, R. (2013). The Role of Communication in Agile Systems Development An Analysis of the State of the Art. *Business & Information Systems Engineering, 5*(5), 338-350.

Humphrey, W. S. (1995). *A discipline for software engineering*: Addison-Wesley Longman Publishing Co., Inc.

Hwang, B.-G., & Ng, W. J. (2013). Project management knowledge and skills for green construction: Overcoming challenges. *International Journal of Project Management, 31*(2), 272-284.

Kan, S. H. (2002). *Metrics and models in software quality engineering*: Addison-Wesley Longman Publishing Co., Inc.

Karhatsu, H., Ikonen, M., Kettunen, P., Fagerholm, F., & Abrahamsson, P. (2010). *Building blocks for self-organizing software development teams a framework model and empirical pilot study*. Paper presented at the 2nd International Conference on Software Technology and Engineering (ICSTE) San Juan, Puerto Rico, USA.

Knaster, R., & Leffingwell, D. (2017). *SAFe 4.0 Distilled: Applying the Scaled Agile Framework for Lean Software and Systems Engineering*. Boston, MA: Addison-Wesley Professional.

Kruchten, P. (2000). *The Rational Unified Process: An Introduction* (2nd ed.). Reading, MA: Addison-Wesley Professional.

Larman, C., & Vodde, B. (2016). *Large-scale scrum: More with LeSS*. Indianapolis, IN: Addison-Wesley Professional.

Lowry, P. B., & Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication, 57*(2), 123-146.

Mao, J.-Y., Lee, J.-N., & Deng, C.-P. (2008). Vendors' perspectives on trust and control in offshore information systems outsourcing. *Information & Management, 45*(7), 482-492.

Maxwell, K. D., & Forselius, P. (2000). Benchmarking software development productivity. *IEEE Software, 17*(1), 80-88.

McLeod, L., Doolin, B., & MacDonell, S. G. (2012). A perspective-based understanding of project success. *Project Management Journal, 43*(5), 68-86.

McLeod, L., & MacDonell, S. G. (2011). Factors that affect software systems development project outcomes: A survey of research. *ACM Computing Surveys, 43*(4), 1-56.

Mishra, D., Mishra, A., & Ostrovska, S. (2012). Impact of physical ambiance on communication, collaboration and coordination in agile software development: An empirical evaluation. *Information and Software Technology, 54*(10), 1067-1078.

Misirli, A. T., & Bener, A. B. (2014). Bayesian networks for evidence-based decision-making in software engineering. *IEEE Transactions on Software Engineering, 40*(6), 533-554.

Misra, S. C., Kumar, V., & Kumar, U. (2009). Identifying some important success factors in adopting agile software development practices. *Journal of Systems and Software, 82*(11), 1869-1890.

Moe, N. B., Aurum, A., & Dyba, T. (2012). Challenges of shared decision-making: A multiple case study of agile software development. *Information and Software Technology, 54*(8), 853-865. doi:DOI 10.1016/j.infsof.2011.11.006

Munns, A., & Bjeirmi, B. F. (1996). The role of project management in achieving project success. *International Journal of Project Management, 14*(2), 81-87.

Murphy, B., Bird, C., Zimmermann, T., Williams, L., Nagappan, N., & Begel, A. (2013). *Have agile techniques been the silver bullet for software development at Microsoft?* Paper presented at the 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement

Nerur, S., Mahapatra, R., & Mangalaraj, G. (2005). Challenges of migrating to agile methodologies. *Communications of the ACM, 48*(5), 72-78.

Papke-Shields, K. E., Beise, C., & Quan, J. (2010). Do project managers practice what they preach, and does it matter to project success? *International Journal of Project Management, 28*(7), 650-662.

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*: Basic Books.

Ramesh, B., Cao, L., & Baskerville, R. (2010). Agile requirements engineering practices and challenges: an empirical study. *Information Systems Journal, 20*(5), 449-480.

Richter, N. F., Cepeda, G., Roldán, J. L., & Ringle, C. M. (2015). European management research using partial least squares structural equation modeling (PLS-SEM). *European Management Journal, 33*(1), 1-3.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. *Boenningstedt: SmartPLS GmbH*.

Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares

path modeling. *Organizational Research Methods, 16*(3), 425-448.
Rönkkö, M., Parkkila, K., & Ylitalo, J. (2012). *Use of Partial Least Squares as a Theory Testing Tool-an Analysis of Information Systems Papers.* Paper presented at the ECIS 2012 Conference.
Rose, K. H. (2013). *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)—Fifth Edition*. Newtown Square, PA: Project Management Institute.
Ryan, S., & O'Connor, R. V. (2013). Acquiring and sharing tacit knowledge in software development teams: An empirical study. *Information and Software Technology, 55*(9), 1614-1624.
Sabherwal, R. (1999). The role of trust in outsourced IS development projects. *Communications of the ACM, 42*(2), 80-86.
Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review, 63*(1), 34-62.
Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research, 69*(10), 3998-4010.
Sarstedt, M., & Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modeling: a comparison of FIMIX-PLS with different data analysis strategies. *Journal of applied statistics, 37*(8), 1299-1318.
Schwaber, K. (2004). *Agile project management with Scrum.* Redmond, WA: Microsoft Press.
Schwaber, K. (2007). *The Enterprise and Scrum*: Microsoft Press.
Senapathi, M., & Srinivasan, A. (2012). Understanding post-adoptive agile usage: An exploratory cross-case analysis. *Journal of Systems and Software, 85*(6), 1255-1268. doi:DOI 10.1016/j.jss.2012.02.025
Sharda, R., Delen, D., & Turban, E. (2016). *Business Intelligence: A Managerial Approach*: Pearson Education.
Siau, K., Long, Y., & Ling, M. (2010). Toward a unified model of information systems development success. *Journal of Database Management (JDM), 21*(1), 80-101.
Siegel, E. (2016). *Predictive analytics: The power to predict who will click, buy, lie, or die*: John Wiley & Sons Incorporated.
Tan, T., Li, Q., Boehm, B., Yang, Y., He, M., & Moazeni, R. (2009, Oct 2009). *Productivity trends in incremental and iterative software development.* Paper presented at the Third International Symposium on Empirical Software Engineering and Measurement, Lake Buena Vista, FL.
Turk, D., France, R., & Rumpe, B. (2005). Assumptions underlying agile software-development processes. *Journal of Database Management, 16*(4), 62-87.
Van Waardenburg, G., & Van Vliet, H. (2013). When agile meets the enterprise. *Information and Software Technology, 55*(12), 2154-2171.
VersionOne. (2016). 11th annual state of agile survey.
Vijayasarathy, L., & Turk, D. (2008). Agile software development: A survey of early adopters. *Journal of Information Technology Management, 19*(2), 1-8.
Vinzi, V. E., Trinchera, L., & Amato, S. (2010). PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement. In *Handbook of partial least squares* (pp. 47-82): Springer.
Vlietland, J., & van Vliet, H. (2015). Towards a governance framework for chains of Scrum teams. *Information and Software Technology, 57*, 52-65.
Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science, 44*(1), 119-134.
Wallace, L., & Keil, M. (2004). Software project risks and their effect on outcomes. *Communications of the ACM, 47*(4), 68-73.

Xia, W., & Lee, G. (2003). Complexity of information systems development projects: conceptualization and measurement development. *Journal of Management Information Systems, 22*(1), 45-83.

Appendix A

Data Warehousing, BI, and Analytics (DW/BIA) Questionnaire

Project Success

1. The project meets or is expected to meet the budgetary estimate.

2. The project meets or is expected to meet the schedule estimate.

3. The project meets or is expected to meet the customer requirements.

4. The project improves or is expected to improve decision-making.

5. The project meets or is expected to meet quality requirements.

Agile values

1. We value individuals and interactions over processes and tools as an important aspect of DW/BIA development.

2. We value working software over comprehensive documentation as an important aspect of DW/BIA development.

3. We value customer collaboration over contract negotiation as an important aspect of DW/BIA development.

4. We value responding to change over following a plan as an important aspect of DW/BIA development.

Plan-Driven Aspects

1. The project has processes to manage scope creep.

2. The project has processes to manage customer expectations.

3. The project has processes to manage commitments such as contracts or promises.

4. The project has processes to manage controls on schedule, cost, and quality.

5. The project has processes to manage risk.

Shared Understanding

1. For reaching a common understanding, IT and business members use dialogue to remove ambiguities.

2. For reaching a common understanding, IT and business members use dialogue to remove domain knowledge gaps.

3. For reaching a common understanding, IT and business members participate in information sharing

sessions.

4. For reaching a common understanding, IT and business members use dialogue to understand the other party's perspective.

Top Management Commitment

1. For the successful completion of the project, the top management provides the leadership.

2. For the successful completion of the project, the top management provides the governance structure.

3. For the successful completion of the project, the top management is willing to invest a great deal of effort beyond that is normally expected.

4. For the successful completion of the project, the top management is committed.

Technological Capability

1. The development team has the ability to employ technological tools to deliver solutions that meet the business requirements.

2. The development team has the ability to employ technological tools to deliver solutions that meet the functional requirements.

3. The development team has the ability to employ technological tools to deliver solutions that meet the non-functional requirements.

4. The development team has the ability to employ technological tools to deliver solutions that meet the architecture requirements.

Complexity

1. During the project completion, difficulties were caused by a large number of pieces during integration.

2. During the project completion, difficulties were caused by the reconciliation of inadequate quality of incoming data.

3. During the project completion, difficulties were caused by the frequent changes in requirements.

4. During the project completion, difficulties were caused by the excess coordination effort among the stakeholders.

Author Biography

Dinesh Batra is a professor in the Department of Information Systems and Business Analytics at Florida International University. Dr. Batra's publications have appeared in Management Science, Journal of MIS, Communications of the ACM, Journal of Database Management, European Journal of Information Systems, Decision Support Systems, Communications of the AIS, International Journal of Human Computer Studies, Data Base for Advances in Information Systems, Information and Management, Requirements Engineering Journal, Information Systems Management, and other journals. He is a co-author of the book Object-Oriented Systems Analysis and Design published by Pearson Prentice-Hall. He served as the first President of the AIS SIG on Systems Analysis and Design (SIGSAND).