The VDC Scorecard: Formulation and Validation

By

Calvin Kam, Devini Senaratna, Brian McKinney, Yao Xiao, & Min Song

CIFE Working Paper #WP135
October 2013, Revised January 2014

STANFORD UNIVERSITY
Abstract

Recent success of virtual design and construction (VDC) can be attributed to the wider adoption of product modeling or building information modeling (BIM) in the AEC industry. Despite this fact, the development of methodologies to assess the maturity of VDC implementation is lagging compared with other innovations in the AEC industry, such as green building assessment frameworks. This is evidenced by the fact that the existing methodologies for assessing the maturity of VDC have been applied to a limited number of projects for short time scales, and that hence the methodologies have been intermittent or static. Most of the methodologies also pay attention to only technical aspects, overlooking social collaboration in assessment. To address this problem, researchers at Stanford University’s Center for Integrated Facility Engineering (CIFE) have drawn from existing research and observations of professional practice to formulate the VDC Scorecard with the goal of making an assessment methodology adaptive, quantifiable, holistic, and practical. The VDC Scorecard assesses the maturity of the VDC implementation of a project across 4 Areas, 10 Divisions, and 56 Measures, and deploys the Confidence Level measured by 7 factors to indicate the accuracy of scores. To keep up with the rapid change of technologies, it aims to build an adaptive scoring system based on evolving industry norms instead of prefixed norms valid for a short period. However, since the industry norms of 4 Areas, 10 Divisions, and 56 Measures could not be compiled at the beginning of the research, the definitions of the percentile of the five tiers of practice (conventional practice, typical practice, advanced practice, best practice, and innovative practice) were drawn from subject matter experts’ opinions. These experts’ percentile for each measure was then calibrated based on the actual sample percentile collected from 108 projects. A correlation test was also done between
individual measures and the overall score, in which measures that were found to correlate poorly were revisited for review and revision. The data sets came from 2 countries in North America, 5 in Europe, 4 in Asia, and 1 in Oceania. They cover 11 facility types, 5 delivery methods, and 5 project phases, making the scorecard holistic. The research process also involved the development of a manual, a web interface, and lite version to facilitate communication and understanding, making the scorecard practical. The initial goal of developing the VDC Scorecard was to make it adaptive, quantifiable, holistic, and practical, and the results and practitioners’ feedback from the research have proven these characteristics.

CE Database subject headings: building information models, assessment, evaluation, project management

1. Introduction

The history of three-dimensional (3D) modeling goes back more than four decades—primitive research on 3D modeling began in the 1960s and 3D modeling migrated to commercial use in the 1970s, largely in the automotive and aerospace industry (Bozdoc 2003). The Architecture, Engineering and Construction (AEC) industry started its 3D modeling research in the 1960s, with wider adoption in the 1980s when cheaper, more powerful personal computers became available. The AEC industry first used 3D models for simple representations of geometric information, but since then 3D models have become more complex and, as of now, utilize
information beyond geometry such as materials, specifications, parametric relationships, and other attributes.

This “digital representation of physical and functional characteristics of a facility,” referred to as Building Information Modeling (BIM) (NIBS 2007), allowed AEC professionals to achieve better building performance, lower costs, shorter schedules, improved communication, higher quality, and improved worker safety. These goals are most readily achieved when project team members choose the technical and social methods involved in using these models specifically to optimize project outcomes. This broader set of considerations leads to Virtual Design and Construction (VDC)—the term which was introduced in 2009 at CIFE (Kunz and Fischer 2009) shortly before the appearance of the term “BIM” in 2002 (Laiserin 2002). The scope of the definition of VDC is broader than that of BIM. The Center for Integrated Facility Engineering (CIFE) defines VDC as the use of multi-disciplinary performance models of construction projects, including their product, organization, and process (POP) models, for business objectives (Kunz and Fischer 2009). While BIM has a tendency to cluster around a product model and the technical aspects of a project, VDC encompasses multi-disciplinary use of POP models and social methods for achieving the business objectives of a project. VDC also stresses the loop between defining objectives and rendering solutions with optimization and automation. Hence while BIM and VDC share similar characteristics, there are subtle additions to VDC in regard to the scope of modeling, the drivers of modeling, and social methods for leveraging those models, making it more comprehensive and holistic than BIM. However, since many entities (e.g., buildingSMART, General Services Administration, American Institute of Architects, Associated General Contractors) and individual projects across the industry set forth their own
definition of BIM, some may argue that BIM also includes these additional characteristics. With this understanding in mind, i.e., with a broader definition of BIM that matches VDC, the VDC Scorecard can be considered as the BIM Scorecard.

The adoption of VDC has been increasing rapidly. As of 2009, more than 80% of major AEC firms in the US had adopted BIM applications (Barista 2009). BIM has also gained dramatic adoption in industry over the past years, increasing from 28% in 2007 to 71% in 2012 (McGraw-Hill Construction 2012). CIFE has also witnessed multi-disciplinary use of those BIM applications and adoption of social methods by which the capabilities of BIM applications can be leveraged. Yet, the AEC industry has not seen a VDC (or BIM) assessment tool that has continually captured or documented the VDC maturity of projects with a consistent framework.

2. Motivation

2.1. Value of Assessment Methodologies

Assessment tools are important since we advance to the stage of science when measurement is involved (Thomson 1889). VDC assessment methodologies can be useful for informed decision-making and to quantify VDC impacts. Management science research findings conclude that good management practices that use innovative tools to track and monitor the management process are correlated with higher productivity, profitability and sales growth rates (Bloom and Van Reenen 2007).
While the implementation of VDC (or BIM) has advanced rapidly, the development of assessment methodologies to measure and facilitate its implementation is lagging, particularly when compared with other areas of performance measurement, such as green building or construction safety. The green building movement can be traced back to the 1970s, when the energy crisis prompted AEC professionals to form groups to lead the movement, such as the energy task force formed by the AIA (Cassidy and Wright 2003). In 1990, the UK’s Building Research Environmental Assessment Method (BREEAM) became the first assessment method for green buildings (Smith et al. 2006). Today every developed country has at least one green building assessment methodology, and the US alone has more than five popular ones: Leadership in Energy and Environmental Design (LEED), Living Building Challenge, Green Globes, Build it Green, National Association of Home Builders National Green Building Standard (NAHB NGBS), and International Green Construction Code (IGCC). This growth has been reciprocated by increased research related to the assessment methodologies for green buildings and, in a broader spectrum, the green building industry.

Developing an assessment methodology for VDC will not only enrich the knowledge for AEC professionals, allowing an accurate assessment of the market, performances, challenges, and trends, but the VDC-related knowledge repository will also bring opportunities for new research and constructive criticism that can be based on a solid and well-substantiated knowledge base. This will create a healthy feedback loop among academia, private and public organizations, and industry groups, leading to the optimization and maximize returns of VDC.

2.2. Limitations of Current Approach
Despite the dearth of assessment methodologies, many professionals and researchers have been conducting theoretical and practical studies to support the industry. At least twenty-two formal papers on the use of VDC on single projects have been published since 1995, and twenty-three notable VDC-related guidelines targeted at the enterprise or industry level are available (Gao 2011). Legal practitioners in collaboration with AEC professionals have also developed supplements to contract documents, such as the glossary and requirements in Appendix A of American Institute of Steel Constructions (AISC) (2010) or the American Institute of Architects (AIA) E202 (AIA 2008) and G202 (AIA 2013) contract exhibits, to help AEC professionals manage and control legal requirements and avoid disputes.

The growth in these areas of study—VDC case studies, guidelines, and standard contracts—has contributed to increasing knowledge about VDC on the individual, team, and trade level. All of these areas of study, however, have their limitations when viewed from a macro industry or global perspective. VDC case studies are unstructured and fragmented because different researchers, individuals, teams, and trades are using their own perspectives or frameworks for their studies, and observation is done only on selective parts of VDC (Gao 2011). Guidelines are also targeted at certain audiences or projects, and they often lack an overarching framework to evaluate the outcome of a project and the usefulness of a given guideline after the execution of a project. Standard contracts, too, are developed on selective parts of VDC. The resulting fragmentation limits the gathered knowledge to small slices of a product, organization, or process of a project. Furthermore the lack of an overarching framework applicable to various
projects across different context makes a fair comparison of different projects difficult, whether on a granular or higher level.

Motivated by this need, a few initiatives have made some progress towards developing a VDC or BIM assessment methodology based on a framework. None of these assessment methodologies are, however, well-recognized or used by the AEC industry practitioners. The reasons for this lack of wide-spread acceptance of these assessment methodologies vary, but they have one or more of the following problems:

1. Lack of objective, metric-based performance measurements
2. Incomplete evaluation of VDC Product, Organization and Process (POP) models
3. Incomplete evaluation of social collaboration in using POP models
4. Application to a limited number of projects for short time scales
5. Lack of investigation into confidence level and demographics of survey respondents
6. Incomplete or unpublished for use
7. Framework or definitions are not intuitive to industry people
8. Lack of industry-ready user interface or access
9. Lack of instructions or manuals for users
10. Limited support from sponsors or contributors
11. Few project cases are used for validation

Since 2009, we have been working towards formulating and validating an effective VDC evaluation methodology for the AEC industry in order to address these problems.
3. Criteria for VDC Evaluation Framework

We noted the importance of creating a standard and comprehensive vocabulary that allows for the objective scoring of VDC and accurate benchmarking of industry practice. In order to accomplish this goal, based on the literature and past experience, we had identified four criteria for a comprehensive VDC evaluation framework. They are:

1) Holistic: Previous frameworks for VDC or BIM, such as National BIM Standards Capability Maturity Model v1.9, tended to focus only on capturing the performances of the creation and process of implementing the technology of a project. In addition to these aspects, a holistic VDC evaluation framework should assess how much the project’s performance is improving through the use of extreme social collaboration addressed in Garcia et al. (2003a) and should be applicable to all phases of a project and account for multiple stakeholders. Furthermore if there are any omissions in capturing the measures of a project, the assessment process needs to inform the accuracy of the score as well. This criterion resolves problems number 2, 3, 4, and 5 stated in section 2.2.

2) Quantifiable: An evaluation framework requires objective and quantifiable measures that can be used for the monitoring and tracking of a project’s progress and VDC maturity. It is established in many industries that "If you can't measure it, you can't manage it" (Jacoby and Luqi 2005). Subjective measures are sometimes un-reliable and difficult to interpret and understand. Therefore, in order to provide accurate and actionable
recommendations for decision making, quantifiable measures, which are inherently easy to monitor and track, are required for an evaluation framework. This criterion resolves problem number 1 stated in section 2.2.

3) Practical: The framework should support decision making among researchers and stakeholders while being reasonably easy to adopt. The evaluation frameworks that are available to the AEC industry today are incomplete, unpublished, regionally applied, research oriented, or nonreciprocal to managers, making them difficult to access by broader practitioners. The evaluation framework needs to be meaningful and actionable for AEC professionals as it depicts a project’s VDC performance over a series of measures in relationship to industry standards. The framework should also leverage previous frameworks, case studies, guidelines, and standard contracts by recommending these precedents as resources where their contents are pertinent to a project’s shortcomings. This criterion resolves problems number 6, 7, 8, 9, and 10 stated in section 2.2.

4) Adaptive: The evaluation framework must be able to adapt to the diversity and evolution of VDC practice. No two AEC projects are the same, thus a framework that is designed specifically around any one type of project will not be practical for much of the industry. Due to the rapidly evolving use of VDC use, the framework must adapt to changing industry practices to avoid obsolescence. This criterion resolves problem number 11 stated in section 2.2.
4. Points of Departure

Researchers and AEC professionals have produced a number of guidelines and research papers on the use of VDC that have influenced the development of the VDC Scorecard. Although few of these documents have overarching contents that cover all of the POP aspects while maintaining practical use in the industry, each of them has contributed in its respective area of knowledge and practice. It is therefore important to build upon their contributions and strengths while considering their limitations and weaknesses in the context of our goals.

Characterization Framework

A thesis from CIFE titled “Characterization Framework to Document and Compare BIM Implementations on Projects” (Gao 2011) describes a framework intended to capture the aspects of VDC in 3 categories, 14 factors, and 71 measures. Incorporated into this framework, the VDC data facilitates sufficient and consistent capture of VDC implementation, and supports across-project comparison. During the development of the framework, it was validated with 40 case projects. This research scope was more focused on capturing the VDC performances from the researchers’ perspective in a particular snapshot. This scope was not to create an active, continuous, repeatable and accessible evaluation interface for practitioners and the framework was not developed to extend the benefits of the framework to practitioners in the industry. In conclusion the Characterization Framework did not provide a practical framework that was adaptive to continuous technological changes.

National BIM Standard - US (NBIMS-US)
NBIMS discusses standardized information related to data modeling for buildings, interoperability, storing and sharing information, and information assurance. It provides a framework, Capability Maturity Model (CMM), which was developed by the University of Florida as a part of the NBIMS (NIBS 2007). The CMM and its more user-friendly version the Interactive CMM (I-CMM) is a framework that is divided into 11 categories and was developed to measure the maturity of BIM implementation. A project can go through an evaluation process with the CMM and be assigned with a rating (Minimum BIM, Certified, Silver, Gold, or Platinum), similar to the LEED system. However, the model is tightly focused on the technical aspects of BIM and shows little signs of assessing the social methods involved in BIM.

**BIM Excellence (BIMe)**

BIMe (BIMe 2014; Succar et al. 2012; Underwood 2009) has thorough structures to define and break down the BIM field, the organizations, and BIM maturity and competency, which is broader than other VDC assessment methodologies developed. But similar to NBIMS, this assessment has gone through limited validation process with actual projects globally. Also the document lacks a quantifiable scoring mechanism that can account for the project outcome objectives (e.g., the target project duration and cost performance) (Kunz and Fischer 2009). For assessing the levels of maturity, BIMe has developed the BIM Maturity Index (BIMMI) after a review of capability maturity models and quality management frameworks—more than 20 of them all together. This BIMMI serves as the foundation for the more detailed BIM Maturity Matrix (BIm³), which has several categories for assessment. The scoring mechanism of BIMe is somewhat passive because, as Underwood (2009) notes, BIm³ intends to minimize the need for
frequent structural changes of the scoring criteria instead of proactively adapting to the growing technologies.

**The BIM Proficiency Matrix**

The BIM proficiency Matrix was developed by Indiana University in 2009. It is designed to understand “the proficiency of a respondent’s skill at working in a BIM environment” and understand the proficiency of BIM in the marketplace (Indiana University 2009). It is a dynamic evaluation tool that is adaptable to project needs. It also has a simplified version as well as an enhanced version. This matrix, however, lacks published knowledge and validation in the research community. The matrix assesses BIM proficiency in 8 categories, but it does not assess project outcome objectives quantitatively and how BIM is supporting them. Also 7 out of 8 categories are for assessing technical aspects of a project, without covering social collaboration.

**BIM QuickScan**

BIM QuickScan was developed in the Netherlands in 2009 to benchmark BIM performance. This evaluation tool was used on over 130 companies and comprised of both quantitative and qualitative assessments of BIM performance (Sebastian and Berlo 2010; Berlo et al. 2012). It comprises of almost 50 multiple-choice online questions divided into 4 chapters, namely: Organization and Management; Mentality and Culture; Information structure and Information flow; Tools and Applications. The key difference between BIM QuickScan and the VDC Scorecard is that the assessment of BIM QuickScan concerns an organization whereas that of the VDC Scorecard concerns a project.
While the VDC Scorecard research investigates the gaps across the current approaches, the intention of the VDC Scorecard research is to collaborate and harmonize constructively with other assessment methodologies above. For instance, the VDC Scorecard is project- rather than organization-oriented. When a problem detected from an assessment is verified to arise from cultural issues within a single organization, the scorecard team can hence instead benchmark the processes of BIM QuickScan since its assessment is organization-oriented.

5. The VDC Scorecard Framework

The VDC Scorecard is an evidence-based methodology, and it evaluates the maturity and performance of VDC in practice based on a rating framework that assesses planning, adoption, technology, and performance. AEC professionals can use the evaluation framework to track and assess VDC performances of their projects. VDC innovation in planning, adoption, technology and performance are defined as 4 Scorecard Areas, whilst 10 sub-Areas are defined as 10 Scorecard Divisions. Under these 10 Scorecard Divisions, there are a total of 56 Scorecard Measures that are evaluated quantitatively or qualitatively (Figure 1). Quantitative measures are those that have a numeric or specifically measurable nature, while qualitative measures are subjective rather than objective. In addition the VDC Scorecard introduces the “Confidence Level” to communicate the degree of certainty of the VDC score, based on the quality of information obtained in the scoring process.
Figure 1: VDC Scorecard Evaluation Framework. The 10 Scorecard Division scores are created using the 56 Scorecard Measures, in turn the 4 Scorecard Area scores are created using the 10 Scorecard Division scores and finally the total VDC score is created using a weighted sum of the 4 Scorecard Area scores.

### 5.1. Scoring System

Drawing from existing precedents and their own research, we developed a percentage based scoring system to quantify the Overall VDC Scorecard scores, Area scores and Division scores. The percentages and tiers inform researchers or professionals as to how the project is performing in relation with the proven practices of the rest of the industry. Thus, the scores have a concrete meaning rather than being an arbitrary value assigned to a performance. This helps AEC
practitioners understand what their score means and gives researchers working on the VDC Scorecard a target against which to base the scoring system.

To further facilitate interpretation of the scores, the percentages were overlaid with the following tiers termed as Maturity Levels of VDC Practice, similar to the LEED Green Building Rating System and the NBIMS CMM and I-CMM. For example the LEED Green Building Rating system gives a score out of 100 points (with 10 additional bonus points), and groups this score into 4 tiers. Similarly the tiers of the VDC Scorecard are 1) Conventional Practice (0%-25%), 2) Typical Practice (25%-50%), 3) Advanced Practice (50%-75%), 4) Best Practice (75%-90%), and 5) Innovative Practice (90%-100%). For instance, saying that a project falls under a conventional practice means that this project is in the 25th percentile among the projects surveyed with the scorecard. The percentiles and the tiers are set through observations of industry practice that are then calibrated using statistical analysis to reflect evolving VDC norms as depicted in the information gathered through the VDC Scorecard’s continued use. Since the norm and what falls under each tier evolve, the scores of a project are coupled with the version of the VDC Scorecard, e.g., 50 percentile score with the VDC Scorecard version 8.

5.2. Overall Score, Area Scores, Division Scores and Process of Scoring

The Overall Score, Area Scores, Division Scores and Confidence Level are presented in scale of 0-100%. The VDC Scorecard’s “overall score” is a measure created using a weighted average of the Area scores corresponding to the 4 Scorecard Areas of Planning, Adoption, Technology and Performance, to quantify the overall VDC performance of an AEC project that utilizes VDC. The
Area scores are four measures created using a weighted average of respective Scorecard Divisions pertaining to a specific Scorecard Area. The Division scores are 10 measures created using a weighted average of Division-related metrics or measurements.

The VDC Scorecard framework and methodology are applicable in all phases of a project lifecycle, from the pre-design through closeout and operations and maintenance phase. The VDC Scorecard methodology can be and has been implemented during an upstream stage when a project is setting-up initial VDC management objectives, or continuously over the lifecycle of a project, or once in an ad-hoc manner at any stage(s) of the projects lifecycle.

5.3. Four Scorecard Areas

The 4 Scorecard Areas—Planning, Adoption, Technology, and Performance—are formulated to represent the themes introduced in Kunz and Fischer (2009), which include the VDC objectives framework, POP, integrated concurrent engineering (ICE), and the VDC maturity model. The Planning Area encourages establishing VDC objectives that are measurable and structured. It also encourages making the VDC objectives public by setting out VDC Guidelines that establish major procedures for using VDC technologies. Then the target values of the objectives are measured quantitatively or qualitatively in the Performance Area. The information modeling of POP is captured through the Technology Area, whereas social methods for successfully adopting technology, such as ICE, are captured through the Adoption Area. Within this framework, a project’s objectives form the initial and most important considerations, and all of the areas are ultimately a means to support these objectives. Both the Technology and the Adoption Areas
serve as drivers that determine whether the objectives enumerated in the Planning Area result in the desired outcomes in the Performance Area.

**VDC Scorecard → Planning Area**

The Planning Area consists of the 3 Scorecard Divisions—Objective, Standard, and Preparation—and these 3 Scorecard Divisions comprise of 13 individual Scorecard Measures (see Figure 1 for the % of the overall score). The planning score is formulated using a weighted sum of the Division scores, and in turn the Division scores are formulated using a weighted sum of the individual Measures.

VDC planning is instrumental in aligning a wide group of stakeholders and identifying the balanced mix of technologies, hardware, software, and training resources needed for the project. The Objective Division evaluates projects on their inclusion of seven categories of VDC. The Standard Division assesses the establishment of VDC or BIM guides as a means of standardizing the implementation. The Preparation Division assesses the extent to which suitable human and capital resources have been allocated and preparations made towards the efficient implementation of VDC. Assessment of these 3 Scorecard Divisions identifies the strengths and weaknesses of VDC planning and enables directed, actionable advice grounded within any human capital or financial constraints. From the 3 Scorecard Divisions, the philosophy and evaluative metrics of the Objective Division are further illustrated here because this Division has the highest weight under the Area and also is the baseline against which the Performance Area is measured.
The setting of project objectives in the Planning section provides critical support for the VDC Scorecard to be practical and quantitative. Quantitative and qualitative objectives are integral to guiding, motivating, and assessing VDC performance. These objectives are the objectives the project team wishes to promote achieving by implementing VDC. Mature targets and metrics help measure and track performance throughout a project’s lifespan and provide feedback on the project’s BIM investment as well as possible areas of improvement. Mature and valid objectives also help prioritize implementation and identify inefficiencies. Even if the established targets are not met, they are still useful in identifying areas of poor performance and informing the correct level objective maturity for future projects. The maturity of objectives implies to what degree the objectives are measureable or have targets and to what degree they cover the seven objective categories as follows.

In the VDC Scorecard, 7 VDC objective categories are created based on the VDC management objectives in Kunz and Fischer (2009). They are identifiable from hundreds of projects that our center and team have observed or participated in. The 7 categories of VDC objectives include:

1. **Communication** – improve quality and frequency of communication between team members. (E.g., % of stakeholders satisfied with the meeting)

2. **Cost** – reduce project costs with better collaboration, analysis, and project solutions enabled through VDC. (E.g., change order rate)

3. **Schedule** – compact the schedule and reduce its uncertainty with better collaborative processes, faster iterations of solutions, and fewer on site conflicts during construction through VDC. (E.g., schedule conformance $\geq 95\%$)
4. **Facility** – leverage VDC to enhance facility performance in areas such as energy use, occupant satisfaction and thermal comfort through better design outcomes that result from greater analysis during the design process. (E.g., energy use—25% better than 2005)

5. **Safety** – reduce risks during the construction and operation of a building by virtually modeling egress, safety, hazards, and simulating construction safety training. (E.g., recordable incident rate)

6. **Delivery** – maximize owner satisfaction by optimizing the project delivery process. (E.g., submittal response latency—less than 5 days)

7. **Management** – Integrate VDC into organizations to improve project management. Involves developing knowledge on how and when such applications should be applied. (E.g., marketing profitability)

**VDC Scorecard → Performance Area**

The Performance Area consists of the 2 Scorecard Divisions—Quantitative and Qualitative—and these 2 Scorecard Divisions comprise of 12 individual Scorecard Measures (See Figure 1 for the % of the overall score). The Qualitative Division assesses non-quantitative objectives, and the Quantitative Division assesses the achievements and monitoring of objectives with numerical benchmarks of performance.

Quantitative measures are weighed more heavily than qualitative ones as shown in Figure 1 to create a more quantifiable evaluation that contributes to the final percentile scoring system. Quantitative assessment is emphasized in the construction of VDC scorecard because of potential
failings of self-assessment—Kruger and Dunning (1999) reported that the poorest performers on examinations tend to grossly overestimate their abilities and Hill and Betz (2005) reports a larger gap between reported and assessed performance when responses have implications on issues of social desirability. As the VDC Scorecard is dependent on responses of professionals who may genuinely want to believe they have mastered advanced methodologies that could make them more valuable to their companies, they could be inclined to overstate performance in self-assessment. Thus, quantitative metrics are weighed more heavily than qualitative ones to ensure objectivity of evaluation. Yet the quantitative input from respondents may not be accurate or lack evidence during the survey, so the Confidence Level in section 5.4 intends to inform this inaccuracy level.

While the most ideal means of taking a quantifiable measurement of the performance of VDC methodologies would be to find their return on investment, many metrics that projects use in defining VDC management objectives cannot be readily converted into an ROI without separate and extensive research. This type of research, e.g., finding an average or norm dollar value of field-generated RFI, is difficult in most of the cases for project teams whose main job is to execute their project. As one of the goals of the VDC Scorecard is to provide a practical tool for decision makers of a project in a timely manner, the VDC Scorecard evaluates the maturity of performance in lieu of ROI together with other performance indicators under Planning, Adoption and Technology Areas.

While quantitative metrics are emphasized, qualitative measures still provides a valuable supplement to evaluations of quantitative performance. Researchers such as Thomas (2001)
reported the reliability of qualitative studies over quantitative ones in certain circumstances. This phenomenon has also been observed at CIFE where practitioner frustration often points to technical difficulties and errors in operations that would otherwise be more difficult to point out with only quantitative metrics. To measure this qualitative aspect of a project, we have mapped the five tiers of the scoring system for the Qualitative Division with BJ Fogg’s Diamond of User Emotion (Fogg 2012). This is a figure in which one axis of the diamond represents the cost (investment) and the other axis represents the benefit (return), each axis in the scale of three ranks (low/medium/high investment or return).

**VDC Scorecard → Adoption Area**

The Adoption Area consists of the 2 Scorecard Divisions—Organization and Process—and these 2 Scorecard Divisions comprise of 18 individual Scorecard Measures (see Figure 1 for the % of the overall score). Proper VDC planning can only be successfully leveraged if the organization and processes adopt the plan with appropriate roles and responsibilities, incentives, and BIM proficiency throughout the project processes. The Adoption Area surveys a project team or enterprise in deploying its human capital to properly support technology plans, and it assesses multi-stakeholder teams with respect to their responsibilities in technology adoption.

With regard to the VDC Adoption Area, the Organization Division measures the level of involvement and proficiency of the stakeholders in a project team; the Process Division assesses the interactions and relationships between stakeholders and their impact on project performance. Both Scorecard Divisions measure progress towards creating integrated and collaborative
processes that can most effectively leverage the multidisciplinary models made possible by VDC technologies, yet they are frequently overlooked in many evaluation frameworks. Even if a project is supported with state-of-the-art tools with a big enough budget, it cannot capitalize on them without proper expertise and interactions between the users of the technology. A project has to secure professionals with the right skill sets and experience for operating the technologies and has to have effective methods for leveraging the human resources.

Although the Scorecard Divisions here are termed “Organization” and “Process,” they are not representative of organization and process models of a project. The technical models are captured in the Technology Area. For instance, a master schedule integrated with a 3D model, which is a combination of a process model and a product model, is surveyed in the Technology Area. The Adoption Area differs from the Technology Area in that the organization and process in the Adoption Area are not technical models but the collaborative, timing, and social aspects of organization and process.

Our experience shows that Integrated Concurrent Engineering (ICE) and Integrated Project Delivery (IPD) foster and streamline VDC. Attributes of them are accredited with higher scores in the scorecard. ICE is used to accelerate the progress of a project while searching for the optimal solution through collocation of different stakeholders, real-time collaboration, and synchronous communications. Originally pioneered by the aerospace and automotive industries in the mid-1990s (e.g., NASA’s Jet Propulsion Laboratory), it was later adapted by CIFE to become an effective method for applying VDC technologies. Chachere et al. (2004) found a set
of factors that enable high level ICE performance, many of which are apparent in the Measures of the Organization Division. These include:

- Readiness of collocation during different phases, phase coverage of each organization, and phase coverage of VDC.
- Quality of a meeting (Garcia et al. 2003b): meeting effectiveness, meeting efficiency, value index and utilization of online voting for agenda.
- Minimized confusion and disruption by new technologies: the frequency of VDC training, the level of training and signs of resistance.

Integrated Project Delivery (IPD) is also an effective method for applying VDC technologies. IPD can be defined as “a project delivery approach that integrates people, systems, business structures and practices into a process ...” (AIA 2007). In practice, this approach helps different organizations to function as a single larger organization. This integration, in turn, fosters an environment that helps data integration, sharing, and interoperability. Hence the social environment created by IPD enables successful adoption of VDC technologies. The signs of IPD are captured under the Process Division.

**VDC Scorecard → Technology Area**

The Technology Area consists of the 3 Scorecard Divisions—Maturity, Coverage, and Integration—and these 3 Scorecard Divisions comprise of 13 individual Scorecard Measures (See Figure 1 for the % of the overall score).
The Technology Area evaluates the technical aspects of VDC applications employed throughout a project. The product, organization, and process models developed with various tools are the subject of evaluation. The three Technology Divisions provide a tiered evaluation of the technology utilized on a project, considering the analyses and models used during design, their information content and level of detail, and how well this information is exchanged with other applications. Maturity is evaluated by categorizing the implemented technologies in a project into 5 levels of implementation, which builds on the 3-level maturity model described by Kunz and Fischer (2009). The 5 levels are: 1) Visualization, 2) Documentation, 3) Model-based Analysis, 4) Integrated Analysis, and 5) Automation & Optimization. Visualization tools aids in understanding a design, component, or process while Documentation tools aid in generating, organizing, and presenting project-related documentation. Model-based analyses are simulations used to model, understand and predict a variety of facility lifecycle issues. Integrated analyses combine multiple analyses and discipline-specific interests into a single analysis process and finally Automation and Optimization involves software and hardware tools that automate design and construction tasks.

Coverage captures both the extent of building elements modeled and the progression of the level of detail over the phases of the project based on the ASTM Uniformat II Classification for Building Elements, and a 5-level classification system (Level of Development) is used: 1) Conceptual, 2) Approximate Geometry, 3) Precise Geometry, 4) Fabrication, and 5) As-Built (AIA 2008). The Integration Division evaluates the interoperability between BIM tools and information loss in BIM exchanges by capturing “the degree of model elements exchanged”, “information loss after model exchange,” and the relationship between different model uses.
5.4. Confidence Level

The CIFE researchers recognize many uncertainties when evaluating the performance of AEC projects, and therefore established Confidence Level measurements for the VDC Scorecard. Apart from the performance measurements, the Confidence Level indicates the accuracy of the scores by taking into account of the sources, completeness of input and frequency of evaluation. By incorporating the Confidence Level to the project scores, the VDC Scorecard also has a structured and consistent way of tracking how the information was collected for a specific project.

Inclusion of Confidence Level measurement provides a more holistic assessment by informing users of the reliability of the assessment. The initial VDC Scorecards measured only the number and level of respondents to the survey as a way of determining the certainty of the score. More measures have been added as this proved insufficient to fully capture certainty. The final list of factors contributing to the Confidence Level is shown below.

Figure 2: Constituents of the Confidence Level
1. Number and Level of Inputs: Survey inputs from multiple managers at an upper level in an organization have a higher Confidence Level than inputs from specialists at a lower level.

2. Multiple Stakeholder Involvement: The stakeholder leading the VDC effort may have favorable views of VDC implementation versus a resistant stakeholder. Hence a higher confident level is given for inputs from multiple stakeholders.

3. Timing and Phase of Engagement: The VDC Scorecard favors projects near completion versus a project in the pre-design stage.

4. Evidence of Documentation: If the interviewees support their claims with evidence of documentation and an independent audit, they can get the maximum Confidence Level in this category. The VDC Scorecard has lowest Confidence Level in this category if no proof is given during the interview.

5. Frequency of VDC Scorecard survey: The VDC Scorecard use is a methodology to track and assess VDC implementation. If the VDC Scorecard is made a part of project analysis on a more frequent basis, the score could be assumed to be more precise.

6. Total duration of the interview per time period: As the VDC Scorecard is currently an interview based survey tool, the more time that can be spent in gathering project and VDC relevant data, the more the Confidence Level would improve.

7. Completeness of the Survey Input Form: The survey input form is exhaustive and the different versions of the VDC Scorecard cater to the fact that not all project teams would be able to spend the required amount of time to answer every question in the survey input form. This ultimately has the largest impact on certainty as it is used as a multiplier
applied to the score resultant of the other confidence factors. The multiplier ranges from 0 to 1 depending on the percentage of questions answered in the VDC Scorecard.

6. Validation of the VDC Scorecard Framework

As of 2012, CIFE researchers and industry collaborators have assessed 108 unique projects and over 150 evaluations consisting of 11 facility types, including medical facilities, offices, laboratories, courthouses, entertainment facilities, and residences, in 13 countries. Using these project data, CIFE researchers for the first time conducted detailed statistical analysis (Kam et al. 2013) in order to comprehensively validate the effectiveness of the VDC Scorecard in evaluating VDC performance of AEC projects. The statistical analysis was used for validating the industry norm or the percentile scoring system assumed when empirical data were unavailable by the subject matter experts, and in order to ensure that the measures do correlate with the VDC performances of a project. Evaluations of the 108 pilot projects demonstrate the VDC Scorecard’s ability to be a holistic, practical, quantitative and adaptive. As of spring 2013 a total of 123 individuals (70 students, 17 researchers, and 36 professionals) have been involved in the formulation, application, and/or validation of the VDC Scorecard.

SME Percentile and Measures

The percentile scoring system captures industry practices for each Measure so that scores can be assigned for projects accordingly. In the initial pilot projects, industry practices were yet to be surveyed. Hence, the pilot projects employed subject matter experts (SMEs), who had 15 years
of experience working with VDC and BIM applications on over 300 projects of different types across the world, to set initial percentile scores. Figure 3 shows the percentile system transitions from its basis on the assessments of industry experts (SMEs) to an assessment scale based on collected data, and its final calibration, using one of the 56 Measures, # of VDC Management Objectives, as an example. As the application of the VDC Scorecard results in more data available to calibrate the scores with, the data will account for an increasingly large portion of the score, a process which is represented by the middle graph. Ultimately, the percentile system will be based entirely on industry assessments.

![Graph showing development of the percentile system](image)

Figure 3: Development of the percentile system. The SME percentile was developed into the calibrated percentile based on actual samples collected.

As data about the industry is collected, the entire scoring system too, will transition to a system that only uses industry data. As of this writing, the percentile system is a calibrated SME percentile system. Using the prevailing 108 cases, a data-driven VDC Score using principal
component analysis was compared with the incumbent expert driven score (Figure 4). As can be observed from the graphs the two methods correlated almost to perfection.

![Comparison of two methods to develop the final VDC Scorecard Score](image)

Figure 4: Comparison of two methods to develop the final VDC Scorecard Score: (1) Expert opinion-based VDC score and (2) Data-driven statistical VDC Scorecard Score.

The surveyed data were also used to refine Measures. Statistical tests used to test correlations are the parametric t-test, chi-squared test, Mann-Whitney test, Kendall’s tau, Spearman’s rank correlation test and Bootstrapping methodologies for validation. These tests and methods were used to determine the correlation and patterns between individual Measures and overall scores (after adjusting for the effect of individual Measure). Measures that are found to correlate poorly (e.g., Measures in Table 1) were revisited for reviewing the design of the questions. Interviewees also made suggestions when they came across questions with nuances or uncertainties, and noticed social or technical methods new in the industry but not addressed or questioned by the scorecard. Through iterations and revisions, the scorecard has developed into the current version, version 8.
Table 1: Least significant correlations (adjusted) of Scorecard Measures with project performance

<table>
<thead>
<tr>
<th>Question</th>
<th>Statistical Significance (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents covered by BEP/VDC Guides</td>
<td>0.726</td>
</tr>
<tr>
<td>Percentage of product elements modeled in 3D</td>
<td>0.1415</td>
</tr>
<tr>
<td>How formalized is VDC among the stakeholders?</td>
<td>0.1399</td>
</tr>
</tbody>
</table>

To validate the VDC Scorecard qualitatively, researchers obtained feedback from industry collaborators and past CIFE researchers. The following statements, among several others, speak to the success of the VDC Scorecard as an assessment tool:

1. I am a project manager with a Company that has utilized BIM on various multifamily, mixed use, and single-family projects. The VDC Scorecard has enabled our Company to continue to improve from project to project by refining the use of BIM throughout the design, construction, and property management processes. As a vertically integrated Company that handles all aspects of development, the use of the VDC Scorecard will enable us to continue to utilize BIM in the most effective and efficient manners for all current and future projects. (Owner/operator/end-user of a facility)

2. Currently there is no good system to measure whether project teams or offices are adhering to Standards. The VDC scorecard has been a good resource for us to define and deliver metrics through a series interviews and checklists. The results accurately described what we were missing in our process and where required further development. This was supported by industry analysis that was already performed by CIFE. (Director of buildingSMART implementation)
Through the iterations and revisions from version 1 to version 8, we have made incremental improvements toward the objective of making the scorecard holistic, practical, quantifiable, and adaptive. The statements of the professionals and the statistical analysis above, together with other data acquired, show these characteristics of the scorecard.

**Holistic, Practical, Quantifiable and Adaptive**

Holistic: The VDC Scorecard has been applied to 11 facility types with feedback from the interviewees for improvement. The interviewees represent project-wide demographics ranging from a chief executive office to junior employees. The latest version of the scorecard is based on the collective feedback from them. The VDC Scorecard is meant to capture the maturity level of technology used in the industry, but it also accounts for social methods for adopting it. Our findings relate to both social collaboration and the maturity of the virtual models of a project. The VDC Scorecard also has been applied to all phases of a project from the pre-design phase to the operation and maintenance phase. Although the histogram of phases is a bell curve, with more projects evaluated during the construction phase, projects in the upstream and downstream phases are also applicable and have been evaluated.

Quantifiable: The VDC maturity of a project is quantifiable against the industry norm. The score of each Measure is based on a histogram and not on assumptions or fixed categorization. The scoring percentile system started off with assumptions made by subject matter experts but has evolved into a calibrated percentile system based on the data acquired from 108 projects. Apart
from providing the scoring system, the VDC Scorecard also encourages project stakeholders to establish goals that are measurable with specific targets. It highly encourages quantitative objectives by rewarding them under Planning and Performance Areas. In addition, it provides examples of metrics and their targets based on previous research in CIFE (Kunz and Fischer 2009), but project stakeholders can also establish their own metrics and targets to incorporate project-specific conditions over the life cycle of a project.

Practical: Lite and full VDC Scorecard evaluations can be completed (respectively) in less than 30 minutes and 90 minutes. These evaluations can further provide immediate quantitative feedback to the project teams with meaningful scores and possible actionable items for decision-making. By providing percentile scores for the 56 Measures in addition to an overall score, the VDC Scorecard informs practitioners where the project stands in relation to the industry norm and best proven practice, and shows areas of improvement that can make the greatest impact to the overall performance of the project. The concept of the Confidence Level provides a “certainty factor” for the result, i.e., a measure of how certain and comprehensive the results of the VDC Scorecard are. The complimentary statements from professionals, noted above, have shown the usefulness in the industry context.

Adaptive: The scorecard achieves flexibility in its measures by allowing project stakeholders to input the objectives and metrics that are most relevant to their projects. The scorecard is also designed to ask for the presence of innovative practice, of which even the interviewers may not be aware. The scorecard has captured and continually reflected the feedback from interviewees from version 1 to version 8. Hence, it achieves fluidity in its scoring system through continuous
refinement and validation on the basis of data obtained through the Scorecard’s application. Continuous statistical analysis on the projects’ scores also helps to identify industry performance trends as well as benchmarks. Based on these trends, the VDC Scorecard is updated to adapt to evolution.

7. Future Research

Interviewer (scorer) interpretations can be subjective. Interviewers, particularly inexperienced ones, can come to significantly different results even when they are looking at the same data for the same project. Ongoing work includes training interviewers to standardize the interview process and improving the Scorecard manual that defines all terms, Measures and inputs. The Confidence Level also suffered in many of the projects. In 63% of projects, respondents were only able to commit to 1-2 person-hours of interviews. To mitigate this, projects often went through multiple evaluations.

While the data set collected by the VDC Scorecard represents one of the more holistic and complete sets, still more data is needed to set the scoring system to be fully based on industry practices and to continue the validation of the Scorecard. Both of these will happen as the Scorecard is used to build a repository of data on VDC. The continued validation will require researchers to ensure that the Scorecard’s Measures are correlated with high overall scores and to ensure that they can discriminate between different levels of VDC performance. The future research will also relate to introducing the trends of the projects under the categories of innovative, best, advanced, typical, and conventional practice in form of case studies.
8. Conclusion

Since 2009, we have formulated the VDC Scorecard as a holistic, quantifiable, practical, and adaptive evaluation framework that allows for objective assessment of VDC in the AEC industry and accurate benchmarking of the industry practice. The VDC Scorecard offers a vocabulary and a scoring system that provide practical feedback to AEC professionals and researchers using objective, quantitative metrics that measure the maturity of VDC implementation and that adapt to changing industry norms.

The VDC Scorecard is an evaluation framework that can comprehensively describe VDC implementation using the overall score, 4 Scorecard Areas, 10 Scorecard Divisions, and 56 Scorecard Measures. The 4 Scorecard Areas are Planning, Adoption, Technology, and Performance. The Planning Area covers the creation of objectives and standards as well as the availability of technological and fiscal resources that will promote the projects’ business goals. The quantitative and qualitative success in achieving these objectives is measured in the Performance Area. The Adoption Area assesses the organizational and procedural aspects of social methods for adopting technology while the Technology Area assesses the product, organization, and process models implemented across five maturity levels.

The Scorecard uses a percentile scale that has been overlaid with a 5 performance tiers for each of the 56 Measures, and supplemented with an assessment of the certainty in the score. The tiers are Conventional Practice, Typical Practice, Advanced Practice, Best Practice, and Innovative
Practice. When combined with the percentile scale, the tiers allow VDC practitioners and researchers to see how they compare relative to the rest of the AEC industry. Subject matter experts set the initial percentile system, but percentile data is continuously updated with additional project information to reflect the state of the industry. The Confidence Level of the score is determined by 7 factors: 1) The number and level of professionals interviewed, 2) the number of stakeholders interviewed, 3) the timing and phase of engagement, 4) evidence of documentation supporting interviewee claims, 5) the frequency of Scorecard use, 6) the total duration of the interviews, and 7) the completeness of the survey input form.

The continuous recalibration and validation of the Scorecard provide the main impetus for ongoing research in the near term future. By the end of 2012, we had used the VDC Scorecard to assess 108 projects, but more data is required for a percentile scoring system based entirely on empirical data. In addition, correlation between individual Measures and overall scores will be assessed, so that less relevant Measures can be revisited for modification. This process creates a positive feedback loop, whereby the Scorecard serves as assessment methodology for AEC professionals, and their data is used to improve the Scorecard. This positive feedback loop will form the basis for sustaining the Scorecard for the indefinite future. As of spring 2014 over 70 students, 17 researchers, and 36 professionals have been involved in the formulation on the VDC Scorecard.

Acknowledgements: The writers would like to thank acknowledge the contribution of the following individuals as former CIFE researchers: Martin Fischer, John Kunz, Brian McKinney, Yao Xiao, Julian Gonsalves, Catherine Aglibert, Ella Sung, Justin Oldfield, Richard Tsai,
Anthony Zara, Rith Yoeun, Michelle Kim, Ali Ghaemmaghami and Linda Brown. The writers would also like to acknowledge the input of the students who enrolled in CEE 212 and CEE 259 at the Department of Civil and Environmental Engineering of Stanford University in 2010, 2011, 2012, and 2013.

References


   PowerPoint Presentation, The Indiana University Architect’s Office, Bloomington, IN.
   ROI.” J. Electronic Commerce Research, 6(1).
   of AEC projects and industry trends.” Working Paper 136, CIFE, Dept. of Civil Engineering,
   Stanford Univ., Stanford, CA.
   recognizing one’s own incompetence lead to inflated self-assessments.” J. Personality and
   Social Psychology, 77(6), 1121-1134.
   implementation suggestions.” Working Paper 97, CIFE, Dept. of Civil Engineering, Stanford
   University, Stanford, CA.
   Construction, New York, NY.
   scope.” National building information modeling standard (NBIMS): version 1.0 – part 1:
   overview, principles, and methodology, NIBS, Washington, DC, 21.
NIBS. (2012). National building information modeling standard—United States™ version 2,
   NIBS, Washington, DC.


Thomas, J. C. (2001). “Qualitative vs. quantitative: Myths of the culture and practical experience.” *Proc., 34th Hawaii Int. Conf. on System Sciences*, the Institute of Electrical and Electronics Engineers (IEEE) Computer Society, Washington, DC.
