

Metrics to Assess Design Guidance

By

Caroline Clevenger and John Haymaker

CIFE Technical Report #TR191 June 2010; updated February 2011

STANFORD UNIVERSITY

COPYRIGHT © 2010, 2011 BY Center for Integrated Facility Engineering

If you would like to contact the authors, please write to:

c/o CIFE, Civil and Environmental Engineering Dept., Stanford University The Jerry Yang & Akiko Yamazaki Environment & Energy Building 473 Via Ortega, Room 292, Mail Code: 4020 Stanford, CA 94305-4020

METRICS TO ASSESS DESIGN GUIDANCE

Heightened sustainability concerns and emerging technologies give building professionals the desire and ability to explore more alternatives for more objectives. As design challenges become more complicated, and as strategies become more advanced, the need and opportunity emerges to measure processes and to compare the guidance afforded. Through literature review and industry observations, we synthesize a comprehensive framework of definitions and metrics. We apply the metrics to an industry case study to illustrate how they help communicate information about challenges, strategies, and explorations present in the domain of energy efficient design. We measure and compare the guidance provided by applying two strategies to one challenge. The ability to measure guidance marks a valuable step for prescribing design process improvement.

Keywords: Design Process, Guidance, Design strategy, Evaluation, Environmental Design

Managing and reducing the environmental impacts of buildings has become a priority of building stakeholders and the architecture, engineering and construction (AEC) community. For example, the American Institute of Architects (AIA) in the 2030 Challenge (AIA, 2007) and the Federal Government in the Energy Independence and Security Act (FEMP, 2007) both call for zero estimated net annual fossil fuel energy consumption for new building designs by the year 2030. However, maximizing energy performance has proven elusive to industry for years because it requires understanding stochastic, dynamic, continuous event-based systems (Bazjanac, 2006). Performance-based design typically embody complex multi-criteria problems that quickly exceed the limits of human cognition and frequently involve trade-offs and interdependences among variables which make it difficult to elicit meaningful design *guidance* (Papamichael & Protzen, 1993). As project teams today are asked to face the daunting task of identifying transcendent, high performing solutions, the ability to evaluate design *strategies* becomes increasingly critical.

Historically, and still today, much of the AEC industry has relied on variously named precedentbased design, experienced-based design or case-based design *strategies* to help resolve design *challenge*s (Watson & Perera, 1997) (Clevenger & Haymaker, 2009). Precedent-based design is a process of creating a new design by combining and/or adapting previously tested design solutions. It benefits from tacit knowledge, and lessons learned. Using precedent to meet building performance objectives, however, has proven to be less than satisfactory with regard to energy efficiency, and little reason exists to assume that it will be effective in addressing the recently proposed, aggressive energy performance goals. Research has shown that professionals generally lack the tacit understanding necessary to guide energy efficient decision-making in a typical design project (Papamichael et al., 1998).

Performance-based strategies involving computer software simulation were introduced with some success in the 1970's (LBNL, 1982). While improving, the tools remain imperfect. Actual energy performance data frequently fails to meet operational design intent for numerous reasons

including complex building science, sub-par construction and/or insufficient operational practices (Clark, 2001; Bazjanac, 2008; Kunz et al., 2009). This research intentionally disregards issues related to the accuracy of energy models and their divergence from actual building performance. It measures the effectiveness of distinct design *strategies* assuming the underlying energy simulation techniques to be sound.

The primary use of energy models in professional practice to date has been for performance verification of individual design alternatives. Of promise, design *strategies* incorporating building information modeling (BIM), parametric modeling and advanced analysis techniques such as optimization and sensitivity analysis are expanding by orders of magnitude the number of *alternatives* it is possible to analyze within a reasonable amount of time. As innovative design *strategies* emerge resulting in new and powerful *explorations*, design teams need a method to assess the *guidance* provided. We define design *guidance* as variation in *exploration* produced by applying different *strategies* to a given *challenge*. Figure 1, graphically represents this relationship by establishing the dimensions of *design process*.

Figure 1: Diagram of *design process* dimensions. Each axis represents a range from low to high levels of advancement, complication, and guidance for *strategy*, *challenge* and *exploration* respectively. Based on these assessments it is possible to evaluate the level of *guidance* afforded.

This research seeks to gain traction in answering the question:

How much *guidance* does a design *strategy* provide?

To answer this question, a designer needs to clearly delineate performance-based *design processes* in terms of the *challenges* faced, the *strategies* applied, and the *exploration* achieved. A comparison across processes will enable an assessment of *guidance*. We use energy performance as the domain of our study. However, the research applies to performance-based *design processes* in general.

1 EXISTING FRAMEWORKS AND METRICS

Design Theory and Design Research are vast fields with application(s) to a broad spectrum of disciplines. We focus on theory most closely related to architectural design processes. Cross (2001) reviews major historical developments in Design Methodology and observes a forty-year cycle in the characterization of the nature of design, oscillating between design as discipline and design as science. In the latest scientific swing, Takeda et al (1990) identify three models for design process: descriptive, cognitive and computable. Under descriptive, Eckert & Clark (2005) identify three classification models: staged based vs. activity-based models, solution-oriented vs. problem oriented literature, abstract vs. procedural vs. analytical approaches. Other research emphasizes the dynamic rather than static nature of design spaces, stating that co-evolution or redefinition of design spaces may, in fact, be the foundation of creativity (Gero, 1996; Maher et al., 1996; Dorst & Cross, 2001). Design Methodology has developed a deep and rich understanding of the process of design. Our contribution is the organization of *design process* into three discrete dimensions: *challenge, strategy and exploration* and the development of metrics for each dimension.

While metrics have generally proven elusive for design processes as a whole (Briand et al., 1994; Bashir & Thompson, 1997), research has successfully developed metrics which address individual *design process* dimensions. For example, (Phadke & Taguchi, 1987) identified signalto-noise ratios in design *variables* as the basis for evaluating the robustness of a design *challenge*. McManus et al. (2007) use the metrics flexibility, robustness, and survivability to evaluate design *strategy*. Simpson et al. (1996) propose design knowledge and design freedom to measure the flexibility of design *exploration,* and (Dorst & Cross, 2001) compare creativity across various student *explorations*.

Additional research exists which begins to evaluate metrics across dimensions. (Cross, 2004) characterizes the *explorations* of outstanding designers as 'solution–' rather than 'problembased.' Shah et al., (2003) proposed the metrics quantity, variety, quality, and novelty to show how well a design *strategy* explores various design *challenges*. (Chang & Ibbs, 1999) identified meaningful indicators of architecture and engineering (A/E) consultants' performance for various design *challenges*. Numerous theoretical mathematical evaluations have been performed to assess algorithm efficiency (i.e.; speed and accuracy) for various *strategies* with regard to certain types of *challenge* (e.g.; local versus absolute maximums). Limited data exists to compare human *explorations* across *strategies* since parallel or redundant *design processes* are rarely performed in the real-world. With the use of Building Information Modeling (BIM) and other electronic information transfers, however, researchers have become increasingly successfully at measuring the flow of information within real-world construction projects (Tribelsky & Sacks, 2010). In conclusion, while existing research addresses various aspects of the dimensions of *challenge*, *strategy* and *exploration,* research lacks complete and full quantification of all three. We note that existing research addresses the dimension of *challenge* and its associated metrics least well.

Terminology currently used for Design Methodology research lacks precision. Love (2002) reviews nearly 400 texts and shows a range of definitions for 'design' or 'design process' that are unique and insufficiently specific. He concludes that these important core concepts are

indeterminate. Design Space, Problem Space, Solution Space, and Trade-space are all terms used in literature. However, 'the set of all possible design options' called 'Design Space' by Shah et al., (2003), is called 'Trade Space' by Ross & Hastings (2005). Conversely, Woodbury and Burrow (2006) state that Design Space is limited to 'designs that are visited in an exploration process,' excluding unexplored options, in apparent disagreement with the previous definitions. A number of frameworks also already exist relating design *variables,* including fuzzy-logic (Ciftcioglu et al., 1998), set-based design (Simpson et al., 1998), and hierarchical systems (Wang & Liu, 2006). The lack of consistency within the literature and its terms across dimensions, however, demonstrates a need for additional research. Striving for clear communication, we begin by precisely defining the terms and relationships intended to explicitly characterize and measure performance-based design. We use italics throughout this paper to indicate specific reference to the proposed component and process dimension definitions.

2 PERFORMANCE-BASED DESIGN DEFINITIONS

In his discussion of Design Research, Dorst (2008) proposes that explanatory frameworks can be used to prescribe improvement to practice. Building upon previous frameworks for design (Akin, 2001; McManus et al., 2007; Chachere & Haymaker, 2011), Figure 2 illustrates our framework of the components, relationships, and spaces in performance-based design. Set notation for each space is given in the left column, while examples of the components in each space are called out in the right column. Sections $2.2 - 2.4$ define these terms which serve as a foundation for our metric definitions presented in Section 3. Section 4 gives real-world examples of these concepts based on an industry case study.

Figure 2: Performance-based Design Framework: Component map for *design process* in Express-G Notation (ISO, 2004). The framework delineates design spaces (left) and illustrates the basic relationships between components (middle). Specific instances of these components are listed (right).

In our framework, performance-based design is an iterative cycle of *objective* identification, *alternatives* generation, *impact* analysis, and *value* assignment to maximize *value*. We do not distinguish a hierarchy among *variables*, nor do we consider uncertainty within our framework.

2.1 COMPONENTS

Here we present the components in reverse order of Figure 2 to emphasize their cumulative nature.

Stakeholder: a party with a stake in the selection of *alternatives*.

Goal: declaration of intended properties of *alternative(s)* (Lamsweerde, 2001).

- **P**reference: weight assigned to a *goal* by a *stakeholder* (Payne et al., 1999; Chachere & Haymaker, 2011).
- **V**ariable: a design choice to be made. A *variable* can be discreet (e.g., number of windows) or continuous (e.g., building length).
- **O**ption: individual *variable* input(s) (e.g., number of windows = $\{1, 2, \text{ or } 3\}$; building length $= 10-20$ meters).

Decision: the selection of an *option* (e.g., a number of windows $= 2$; building length $= 12.75$ meters)

Alternative: a combination of *decisions* about *options.*

Constraint: limit placed on *variable*.

Requirement*:* limit placed on *impacts.*

Impact: *alternative's* estimated performance according to a specified *goal*. Estimates range from relatively quick and simple to elaborate and detailed and may or may not be easily quantifiable (Earl et al., 2005).

Value: net performance of an *alternative* relative to *preferences*, *goals* and *constraints* (see Equation 1).

2.2 DESIGN SPACES

Building on our components, we define the following spaces illustrated in the left column of Figure 2.

- *Objective Space* { S, G, P, C }: Set of *stakeholders*, *goals*, *preferences* and *constraints.*. These individual components are inter-related since weights and acceptable ranges of performance can never be completely separated (Earl et al., 2005).
- *Alternative Space* { A, uA }: All feasible *alternative*s for a given *challenge*, including explored and unexplored *alternative*s (Tate & Nordlund, 1998). The space is sufficiently vast that it can be thought of effectively unbounded relative to designer's time and reasoning ability (Kotonya & Sommerville, 1997).
- *Impact Space* { I, R }: All analyzed *impacts* for *alternative*s relative to *goals* and determined to be acceptable or unacceptable according to *requirements*.
- *Value Space* { V }: *Values* generated during an exploration. *Value* is a function of an *alternative*'s *impact* and *stakeholder preference* relative to project *goals*.

In addition to these explicit delineations is the implicit frame of the design space. Most similar to "problem space" as defined by others (Dorst & Cross, 2001), we acknowledge that our design space assumes which *variables* or *goals* to include. We limit our research to *decisions* within a design space, and do not include metrics for evaluating the frame of that space.

2.3 PROCESS DIMENSIONS

Based on our defined components and spaces we provide the additional terms to form the dimensions of *design process*.

Challenge: a set of *decisions* to be made ranging from simple to complicated. *Strategy*: a procedure to generate *decisions* ranging from none to advanced. *Exploration*: a history of *decisions* made ranging from random to guided. *Design process*: implementation of a *strategy* to a *challenge* resulting in an *exploration*. *Guidance*: variation in *exploration* produced by applying different *strategies* to a given *challenge*.

3 MEASURING DESIGN PROCESS

We use our defined components, spaces and dimensions to organize and develop our *design process* metrics. Most metrics are normalized from 0 to 1 and, with a few noted exceptions, the higher numbers are generally considered better.

3.1 QUESTIONS MOTIVATING METRICS

The following questions motivate our metrics. Grounded in literature, these questions are organized according to dimension and span performance-based design spaces. In the next section we individually address each of these questions by developing a corresponding numeric measure.

3.1.1 DESIGN PROCESS CHALLENGE

1) How many *objectives* are included in the *challenge* and how clearly are they defined? Designers need to assess the quantity and quality of project *objectives* (Chachere & Haymaker, 2011).

2) To what extent do *objectives* interact? Other researchers have noted that performance *goals* can be in competition (Ross, 2003; McManus et al., 2007). Designers need to understand the extent to which trade-offs exist when assessing the complexity of a *challenge*.

3) To what extent do *decisions* interact? Building science is not a system of independent variables to be sub-optimized. (Deru & Torcellini, 2004; Wang & Liu, 2006; Bazjanac, 2008). Designers need a measure of the interactive effects between *variables* when assessing *challenge*.

4*)* What is the relative *impact* of each *decision*? Research has shown the important role of screening and sensitivity analyses (Kleijnen, 1997.) Designers need a measure of the extent to which the *impact* caused by any one or pair of *variables* dominates *value*.

3.1.2. DESIGN PROCESS STRATEGY

5) Of the *goals* identified, what *goals* does the design *strategy* consider? Performance *goals* are fundamental to performance-based design, and previous research lays the groundwork for defining and assessing the completeness of the *goals* analyzed (Gero, 1990; Ross, 2003; Edvardsson & Hansson, 2005; Chachere & Haymaker 2011).

6) What *alternatives* does the design *strategy* consider? Discrete *alternative*s have been long considered the building-blocks of design (Gero, 1990; Smith & Eppinger,1997). Emerging generative and parametric modeling techniques test the boundaries of "discrete" design *alternative*s (Gane & Haymaker, 2007; Hudson, 2009). Research predominantly supports the hypothesis that generating more *alternative*s increase the chance of high performance (Akin, 2001; Ïpek et al., 2006). Designers need to understand the size and substance of the *alternative space*.

7) How diverse are the investigated *alternatives?* Many researchers have written about the role of creativity in design (Akin & Lin, 1995; Gero, 1996; Dorst & Cross, 2001; Shah et al., 2003). Designers need to assess the diversity of combinations of *options* used to generate *alternative*s in an *exploration*.

3.1.3. DESIGN PROCESS EXPLORATION

8) What is the average performance of *alternatives* generated? Common metrics in descriptive statistics include mean and mode.

9) What is the range of performance of *alternatives* generated? A common metric in descriptive statistics is standard deviation to measure variability within a given data set.

10) How many *alternatives* are generated before best *value* is achieved? Other researchers have studied iterations as well as process efficiency to understand how and how quickly a *strategy* will converge on an solution(s) (Smith & Eppinger, 1997; Wang & Liu, 2006, Chen et al., 2008).

11) What is the best performing *alternative* generated? A common metric in descriptive statistics is maximum *value*. Research in set-based design and pareto-fronts also provides the possibility of multiple optimums in design (Simpson et al, 1998; Ross & Hastings, 2005).

Collectively these questions illuminate information that can help designers to understand a *design process.* In the next section, we use our framework to develop metrics for these questions. We then test these metrics by comparing the *guidance* provided by two different *strategies* in a real-world case study.

3.2 DESIGN PROCESS METRICS

Table 1 defines the specific terms we use to develop metrics that can help to numerically characterize the dimensions of *design process*. In certain instances a complete analysis of the *alternative space* and *value space* is required to evaluate the individual terms.

Table 1: Design Process Terms.

 p_1, \ldots, p_G , *preference* relative to each *goal* analyzed.

 i_{11}, \ldots, i_{AG} , impact of individual *alternatives* relative to *goals* analyzed.

t, total time required to generate and analyze all *options*.

c, the number of *constraints*.

I, importance, the ranked (% of 100) *impact* of a *variable* (or variable pair) on *value*.

IAVG, average rank of *impact* for all *variables*.

IMEDIAN, median rank of *impact* for all *variables*.

IHIGH, rank of *variable* with the highest *impact*.

ItheorecticalHIGH, the highest percentage rank possible in a series, given the median rank of *impact* over all *variables*.

vA, *value* of an *alternative*, the aggregate *impact* of an *alternative* weighted according to *stakeholder preference*.

V, the set of *alternative*s generated with acceptable *impacts*.

Using the terms listed in Table 1, we develop the following metrics to measure *design process*.

OBJECTIVE SPACE SIZE, OSS = {Ga}

OSS is the number of *goals* analyzed by a given *strategy*. This metric is a count, and is not normalized.

For example, if an energy simulation software tool is capable of analyzing energy usage, thermal performance as well as first cost (LBNL, 2008), $OSS = 3$.

ALTERNATIVE SPACE INTERDEPENDENCE, ASI = $\frac{\textbf{n}_{\text{interact}}}{\binom{n}{2}}$

ASI is the number of first order interactions among *variables* divided by the number of *variable* pairs. A high ASI (0 to 1) indicates a higher number of interactions occurring among *variables*. A high ASI contributes to the level of complication of a *challenge*.

In this example, we illustrate interdependence visually. A-symmetry about the X-Y diagonal indicates that an interaction is occurring among *variables*. Visual inspection of Figure 3 demonstrates interdependence between Window Type and HVAC Efficiency (left), HVAC Efficiency and Roof Insulation (center), but no significant interdependence between Window Type and Roof Insulation (right).

Figure 3: *Value* (NPV) as a function of combinations of Window Type, HVAC Efficiency, and Roof Insulation *variables*. The asymmetry of the first two graphs shows two interactions of the first order among the three *variables*.

From the data shown in Figure 3, ASI = $2/3$ = .66.

IMPACT SPACE COMPLEXITY, ISC = $n_{\text{trade-offs}}/n$

ISC is the number of *variables* that result in performance trade-offs (divergent *impacts*) divided by total number of *variables* considered. ISC represents the percent of *variables* for which *goals* are competing. A high ISC (0 to 1) contributes to the level of complication of a *challenge*. In the case where only one *goal* is assessed, ISC, by definition equals zero.

For example, consider the case where 3 *variables* (HVAC Efficiency, Window Type and Exterior Shading) are evaluated relative to the *goals* to minimize energy usage, and minimize first cost. Both HVAC Efficiency and Window Type show competing *impacts*- higher first costs resulting in lower energy usage. However, for Exterior Shading, the first cost increase of the Exterior Shading is offset by cost savings resulting from a downsized HVAC system. In the case of Exterior Shading *impacts* are not competing and the *option* with the lower first cost also has lower energy usage. In this case $ISC = 2 / 3 = .667$.

VALUE SPACE DOMINANCE, VSD =
$$
\frac{I_{\text{AVG}} - I_{\text{MEDIAN}}}{\left(\frac{100}{\text{Nimportance}}\right)} * \frac{I_{\text{HIGH}}}{I_{\text{theoretical-HIGH}}}
$$

VSD is the extent to which *value* is dominated by individual or combinations of *variables*. The metric features the terms average, median, and high rank of variable *impact*. It is a function of the theoretical high rank, over the median rank. A high VSD (0 to 1) indicates that the *value space* is highly dominated and suggests that the *challenge* is not complicated.

We demonstrate VSD using a simple, but extreme example. Consider two cases where *variables* are ranked in terms of their potential effect on *value*. Figure 4, Series 1 represents minimal dominance, Figure 4, Series 2, represents maximum dominance.

Figure 4: Diagrams depicting minimum (left) and maximum (right) dominance among five *variables*. High dominance indicates a high correlation between optimization of a single *variable* and maximum *value*.

Numerically these series have values: *Series 1: 20,20,20,20,20 Series 2: 1,1,1,1,96*

We add a third, less extreme series for illustrative purposes. *Series 3: 5, 10, 15, 25, 45*

In all cases, the numbers in the series sum to 100 since the numbers each represent a percentage impact. Here we calculate the VSD for the three series showing Series 1 being the least dominated, Series 2 the most, and Series 3 partially dominated:

$$
VSD_{series1} = \frac{20 - 20}{\left(\frac{100}{s}\right)} * \frac{20}{20} = 0
$$

$$
VSD_{series2} = \frac{20 - 1}{\left(\frac{100}{s}\right)} * \frac{96}{96} = .95
$$

$$
VSDseries3 = \frac{20 - 15}{\left(\frac{100}{5}\right)} * \frac{45}{55} = .20
$$

OBJECTIVE SPACE QUALITY, OSQ = Ga / G

OSQ is the ratio of the number of *goals* analyzed to the number of *goals* identified. It demonstrates the extent to which (0 to 1) the *strategy* addresses project *goals*.

If, for example, in addition to energy usage, thermal performance and first cost, acoustic performance is important, then for a *strategy* relying exclusively on energy simulation software has an OSQ = 3 / 4 because acoustic *impact* is not assessed.

ALTERNATIVE SPACE SAMPLING, ASS = A / AS ~ A / (A + UA)

ASS is the number of *alternative*s generated divided by the number of *alternative*s required for "significant sampling" of the *alternative space*. It demonstrates the extent to which a *strategy*'s sampling is statistically significant. Significant sampling can be determined using standard mathematical calculations for a statistical "sample size." While such analysis is non-trivial, the mathematical algorithms addressing such anomalies falls outside scope of this research. When the statistically significant sample size is unknown, the total number of possible *alternative*s is used.

If, for example, A_S is unknown, but the *alternative* Space includes 1000 feasible *alternatives*, yet only four *alternatives* are analyzed, then $ASS = 4 / (4 + 996) = .004$

ALTERNATIVE SPACE FLEXIBILITY, ASF = $(\Delta o_{A i A j} /) / n$

ASF is the average number of *option* changes between any two *alternative*s divided by the number of *variables*. ASF measures the level of decision variation in a given *exploration*. ASF is calculated by taking every pair of *alternative*s in a *design process* and averaging how many *variables* have differing *options* between all pairs. Because ASF averages across every combination of *alternative*, sequence of *exploration* becomes immaterial. The metric represents the breadth or diversity of an *exploration*, regardless of sequence.

For example, the following *exploration* consists of three *alternative*s, each including three Variables.

Alternative 1: Low Efficiency HVAC, Single Pane Windows, Low Roof Insulation *Alternative* 2: Low Efficiency HVAC, Single Pane Windows, High Roof Insulation *Alternative* 3: Low Efficiency HVAC, Double Pane-LowE Windows, High Roof Insulation *Alternative* 1 to Alternative 2: 1 *option* change *Alternative* 1 to Alternative 3: 2 *option* changes *Alternative* 2 to Alternative 3: 1 *option* change

 $ASF = ((1+2+1)/3)/3 = .444$

VALUE SPACE AVERAGE, VSA = ^V

VSA is the mean *value* for the set of *alternative*s analyzed. It characterizes the average *alternative* generated in an *exploration*.

For example,

 $NPV_{Alternatively1} = 25 $NPV_{Alternative2} = 32 $NPV_{Alternative3} = 30

 $VSA = 29

VALUE SPACE RANGE, VSR = STDEV(v_I **)**

VSR is the standard deviation of all *values* for the set of *alternative*s analyzed. It characterizes the dispersion of *alternative*s generated in an *exploration*.

For example,

 $NPV_{Alternativelyel} = 25 $NPV_{Alternative2} = 32 $NPV_{Alternative3} = 30

 $VSR = 3.6

VALUE SPACE ITERATIONS,VSI= Number of *alternatives* **generated prior to achieving maximum** *value*

VSI is the number of *alternative*s generated before the highest *value* is reached. It characterizes the efficiency of an *exploration* and is to be minimized.

For example,

 $NPV_{Alternative1} = 25 $NPV_{Alternative2} = 32 $NPV_{\text{Alternative3}} = 30

 $VSI = 2$

VALUE SPACE MAXIMUM, VSM = MAX(vi)

VSM is the highest *value* of all *alternative*s generated. It identifies the maximum performance achieved in an *exploration*.

For example, $NPV_{Alternatively1} = 25 $NPV_{Alternative2} = 32 $NPV_{\text{Alternative3}} = 30

 $VSM = 32

In the next section we use these metrics to measure and compare *design processes* in real-world, industry case studies.

4 INDUSTRY CASE STUDIES

To test and illustrate our metrics, we applied them to an industry case study. The first documents a professional energy analysis performed in 2006 during schematic design of a 338,880sf Federal Building with 10 floors and a parking sub-basement sited in a mixed (hot in summer, cold in winter), dry climate at an elevation of 4220ft. At the beginning of the project, the client set an annual energy usage target of 55 kBtu/sf/yr as a *requirement* for the project. Additional *goals* were low first cost, and pleasing aesthetics. The mechanical engineer on the project used DOE-2 (LBNL, 1982) to simulate energy performance. A total of 13 energy simulation runs were generated during 4 rounds of energy modeling. Figure 5 represents the *alternatives* modeled and associated annual energy savings (kBTU/sf/yr) estimates generated by professional energy modelers and delivered to the project team in several reports in table or written narrative format.

The case study represents a process which involves multiple parties in a collective decision making process, an arrangement which has been shown to protract decision making (Tribelsky & Sacks, 2010). The example is, nevertheless, relevant to our research because collaborative design, distinct from concurrent or cooperative design, has the distinguishing characteristics of shared *objective*(s) (Ostergaard & Summers, 2009). As such professional analysis using

collaborative design can be meaningfully compared with advanced analysis, since each is reliant singular rather than multi-objective decision-making process.

4.1. PROFESSIONAL DESIGN PROCESS

Results from the 13 energy simulations, generated over a 27 month period are summarized in Figure 5. The black line shows estimated annual energy savings (kBTU/sf/yr) for individual whole building simulations during Schematic Design. The *strategy* for generating and analyzing the *alternative*s can primarily be characterized as performance verification: performance "pointdata" is provided as individual design *alternative*s are generated for the primary purpose of verifying performance relative to the performance *goal(s)* as well as the previous runs.

Alternatives Generated by Professionals through Time

Figure 5: Graphical representation of a professional energy modeling during the design process. *Variables* are listed on the right. *Alternatives* are shown as vertical stacks of specific combinations of *options* (represented in greyscale.) Changes in *options* for each *alternative* can be observed through horizontal color change. Estimated energy saving is depicted with a solid dark grey line. The dashed light grey line shows target energy savings. The figure suggests that the professional energy modeling performed using this *strategy* supported a slow, disjointed, unsystematic *exploration*.

Additional detail regarding the *variables* and *options* analyzed in the professional *exploration* case is provided in Table 3 in the appendix. Energy savings are calculated relative to a professionally estimated baseline, which assumes the minimum inputs necessary for prescriptive code compliance.

4.2. ADVANCED DESIGN PROCESS

An emerging technique to support the development of multidisciplinary design and analysis strategies is Process Integration and Design Optimization (PIDO) tools such as those provided by Phoenix Integration (Phoenix, 2004). PIDO integrates 3-dimensional parametric representations and analysis packages to facilitate the rapid and systematic iteration and analysis of geometric and non-geometric variables. Recent work applied this technique to the energy efficiency domain in AEC (Welle & Haymaker, 2011). In our research, we used PIDO to support application of an advanced strategy to the industry case study. We implemented a design of experiment to explore

the *variables* described in the professional exploration to estimate annual energy savings (Figure 6), and to evaluate trade-offs between first cost and energy consumption. While it is possible to easily expand this method to explore many more *variables* and/or *options*, we chose to scale the *exploration* using our advanced *strategy* to be similar to the scale of the professional analysis to facilitate comparison. For additional detail regarding the *variables* and *options* analyzed in the professional exploration case, see Table 4 in the appendix. Energy savings are calculated relative to the professionally estimated baseline, which assumes the minimum inputs necessary for prescriptive code compliance.

Figure 6: Graphical representation of advanced energy modeling process using PIDO. 1280 *alternatives* are represented as grey diamonds. Each *alternative* changes a single *option*. Y-axis shows the estimated energy savings of a given *alternative*. The 13 energy performance estimates generated by professional modelers (Figure 5) are overlaid in black. The dashed black line shows target energy savings. The figure contrasts the more systematic full *design process* supported by the advanced strategy to professional practice. It demonstrates that a significant number of *alternatives*, unanalyzed in professional energy modeling, have superior *value*, and exceed target performance.

To compare the *strategy* used to support professional practice today to emerging advanced *strategies* being studied in research, we applied our metrics to two processes using each *strategy*. Calculations, assessments and comparisons of the two processes are presented in Table 2.

Table 2: Metrics evaluated comparing traditional professional design processes to the advanced process We developed using PIDO.

* see appendix, Figure 8, ** see appendix, Figure 9, *** see appendix, Figure 10

Challenge metrics for the two case studies are presumed to be closely aligned. In traditional energy modeling, however, neither statistical sampling nor full analysis is performed and direct assessment of *challenge* metrics is not possible. We assume the *challenge* metrics assessed using the advanced *strategy* apply to both case studies since the *challenge*s contain only minor modeling differences due differences necessitated by modeling. The advanced *strategy* reveals that *value* in the case study is highly dominated (VSD = .68) by one decision, window area (see Appendix, Figure 8). In the traditional process, the designers displayed a strong *preference* for an all-glass exterior having qualitative but not quantitative knowledge of the extent of its dominance. The high impact of the decision regarding window area is observable in Figure 6, where estimated energy savings dramatically drops between *Alternative* 3 and *Alternative* 4 due to an increase in window area. Interestingly, in traditional practice the designers never revisited the decision regarding "window area," but maintained the 95% exterior glass *option* for all remaining *alternative*s explored. Alternative Space Interdependence (ASI) from the advanced *strategy* reveals that nearly half of the *variables* modeled have some level of dependency. This result is not surprising since building geometry (e.g., square or rectangular) was a design *variable* that affected nearly every other *variable* modeled (e.g., percentages of window or

shading). Finally, Impact Space Complexity (ISC) in the advanced *strategy* shows relatively little competition between first cost and energy savings. This result is unintuitive and is a function of the "self-sizing" HVAC currently modeled. In other words, although energy efficiency measures may have a higher first cost, these are partially offset by the cost savings that result from a smaller HVAC system.

Strategy metrics are similar in Objective Space Quality (OSQ). Both *strategies* quantify energy savings and first cost, but do not directly assess aesthetics. Such assessment is left to designer judgment. Alternative Space Sampling (ASS) score for the advanced *strategy* is orders of magnitude better than the traditional *strategy*. By scripting and queuing the execution of model simulation runs, the advanced *strategy* performs full analysis (ASS = 1) for all feasible *options* of 9 *variables* in a fraction of the time (4 hours versus 2.3 mo.) compared to the traditional *design process*, which relies upon manual model revision to execute a total of 13 runs. Alternative Space Flexibility (ASF), using the traditional *strategy*, however, is higher. On average, each *alternative* differs by three *options* when manually selected while only one *variable* at a time is changed according to the script of the advanced *strategy*.

Exploration metrics for the advanced process show improved maximum and average *value* generated. In the case of advanced analysis, we assume a designer would merely select the top performer identified. We were able to perform full analysis for our case study since the total number of runs required was relatively small. Other research has addressed much more complicated *challenges* with nearly exponential number of runs. Nevertheless, our case is informative since the level of *challenge* mimics the one actually addressed through professional energy modeling in our real-world case study. In the future, as more advanced modeling capabilities come on-line, we predict that statistically significant sample sizes will frequently become the norm and that the metric, Value Space Iteration (VSI), will likely become relatively insignificant and will primarily depend on computing power.

While multiple metrics support the characterization of each of *design process* dimension, it is possible to crudely graph these relationships by simply summing the metrics without weights. Figure 7 summarizes our findings graphically. Future research is necessary to refine and calibrate these dimensions and their relationships.

Figure 7: Diagram of the *guidance* provided by two different *strategies* representing professional energy modeling and advanced energy modeling applied to the same *challenge*, the design of a federal office building.

Assessment of the metrics suggest that advanced design *strategy* tools being developed by researchers provide better *guidance* than the traditional energy analysis being performed in industry today based on higher Value Space Average (VSA), Value Space Range (VSR) and Value Space Maximum (VSM) (Table 2). Our ability to apply the framework and metrics to test cases is evidence for the claim that they help clarify both relative and absolute *design process* performance assessment.

5 CONCLUSION

In the face of expanding objectives and alternatives, professionals need to choose *design strategies* that efficiently generate high performing *alternative*s. The use of precedent and pointbased analysis *strategies* has proven less than satisfactory in addressing energy efficiency. Significant opportunity exists for advanced *strategies* to provide designers better *guidance* that results in more effective *exploration*s. To realize this potential, designers need a language to compare and evaluate the ability of *strategies* to meet their *challenge*s. Literature review provides a foundation, but not a complete basis for such comparison.

In this paper, we define *design process* to consist of three dimensions: *challenge*, *strategy* and *exploration*. We develop a framework to precisely define the components and spaces of performance-based design. We synthesize a set of metrics to consistently measure all dimensions of *design process.* Finally, we demonstrate the power of the framework and metrics by applying them to the application of two distinct *strategies* on a challenge. We observe that the framework and metrics facilitate comparison and illuminate differences in *design processes*.

Strengths of the metrics include the ability to assess differences in *challenge*s not previously quantified using traditional point-based *design processes*. Alternative Space Flexibility (ASF) is potentially the most important and controversial metric. One interpretation of ASF is as a proxy for creativity. In our case studies, the metric shows full analysis to be the least creative *strategy*. Researchers have long recognized the antagonism between creativity and systematic search and the link between creativity and break-through performance. (Gero,1996; Dorst & Cross, 2001; Shah et al., 2003). Here, we recognize that creativity exists on at least two levels: within set bounds of a decision frame, and beyond (re-formulated) project boundaries. Our ASF metric currently addresses the lesser level of creativity within the bounds of established project *constraints*. The higher level of creativity is not addressed. Similar to the rationale for much computer-assisted design, however, we propose that relieving designers of iterative tasks and examining more *alternative*s and *objectives* potentially enables them to be more creative.

We encountered several areas where improvement and future research is needed. Certainly, full analysis in all but the simplest *challenge*s is not possible in building design. We anticipate that advance *strategies* in real-world applications will rely on sophisticated sampling techniques, modeling simplifications, or precedent-based knowledge bases for *alternative* and/or *objective* formulation. Alternative Space Sampling (ASS), measures the degree to which the number of *alternative*s generated is a representative, statistical sampling of *alternative space*, but says nothing of the distribution of this sampling. Finally, debate remains surrounding the role and potential supremacy of Value Space Maximum (VSM) as a design *exploration* metric. Should a *process* that produces the highest VSM be considered the best regardless of other *exploration* metrics, such as Value Space Average (VSA) generated? In general, the relative weight and relationships of all of the metrics merits further clarification and research.

The strength of these metrics is that they begin to address the eleven questions outlined in Section 3.1, and provide quantitative measure of the three dimensions of *design process*. While we provide evidence of their power using a *challenge* based on a real-world case study related to building energy performance, the findings of this research are not limited to the field of energy efficiency. Future work will expand the application of these metrics to evaluate and compare more and, more advanced and diverse *challenges*, *strategies* and *explorations* in theoretical and real-world design. The metrics enable comparison both within and across dimensions, perhaps indicating which *strategies* are best suited for which *challenges* (Clevenger et al, 2010)*.* In this case, the concept of "process cost" for a *strategy* will need to be explicitly addressed to enable designers to select the best *strategies* for a particular design (Clevenger & Haymaker, 2010). In general, this research allows designers to better evaluate existing and emerging *design processes* and, potentially, to prescribe improvement to practice.

References

- American Institute of Architects (AIA), (2007). National Association of Counties Adopts AIA Challenge of Carbon Neutral Public Buildings by 2030.
- Akin, Ö. (2001). Variants in design cognition. In C. Eastman, M. McCracken & W. Newstetter(Eds.), Design knowing and learning: Cognition in design education (pp. 105- 124). Amsterdam: Elsevier Science.
- Akin, Ö., & Lin, C. (1995). Design protocol data and novel design decisions. *Design Studies, 16*(2), 211-236.
- Bashir, H. A., & Thompson, V. (1997). Metrics for design projects: a review. *Design Studies, 20*(3), 163-277.
- Bazjanac, V. (2006). *Building energy performance simulation presentation*.
- Bazjanac, V. (2008). IFC BIM-based methodology for semi-automated building energy performance simulation. *Lawrence Berkley National Laboratory (LBNL), 919E*.
- Briand, L., Morasca, S., & Basili, V. (1994). Defining and validating high-level design metrics.
- Chachere, J., & Haymaker, J. (2011). "Framework for measuring rationale clarity of AEC design decisions." ASCE Journal of Architectural Engineering, [10.1061/\(ASCE\)AE.1943-](http://dx.doi.org/10.1061/%28ASCE%29AE.1943-5568.0000036) [5568.0000036](http://dx.doi.org/10.1061/%28ASCE%29AE.1943-5568.0000036) (February 8, 2011).
- Chang, A., & Ibbs, C. W. (1999). "Designing levels for A/E consultant performance measures." Project Management Journal, 30(4), 42-55.
- Chen, C. H., He, D.,, Fu, M. C., & Lee, L. H., (2008). Efficient simulation budget allocation for selecting an optimal subset. *INFORMS Journal on Computing* accepted for publication.
- Ciftcioglu, O., Sariyildiz, S., & van der Veer, P. (1998). Integrated building design decision support with fuzzy logic. *Computational Mechanics Inc.*, 11-14.
- Clarke, J. A. (2001). Energy simulation in building design, Butterworth-Heinemann.
- Clevenger, C., Haymaker, J., (2009). [Framework and Metrics for Assessing the Guidance of](http://mycahs.colostate.edu/caroline.clevenger/documents/ICED09-449-SUBMITTED.pdf) [Design Processes,](http://mycahs.colostate.edu/caroline.clevenger/documents/ICED09-449-SUBMITTED.pdf) *The 17th International Conference on Engineering Design*, Stanford, California.
- Clevenger, C., Haymaker, J., Ehrich, A. (2010). Design Exploration Assessment Methodology: Testing the Guidance of Design Processes*,* Available from:

http://cife.stanford.edu/online.publications/TR192.pdf [Accessed June, 2010].

- Clevenger, C., Haymaker, J., (2010). Calculating the Value of Strategy to Challenge, Available from:<http://cife.stanford.edu/online.publications/TR193.pdf> [Accessed June, 2010]
- Cross, N. (2001). Designerly Ways of Knowing: Design Discipline versus Design Science, Design Issues, Vol. 17, No. 3, pp. 49-55.
- Cross, N. (2004). Expertise in design: an overview. *Design Studies, 25*(5), 427-441.
- Deru, M., & Torcellini, P. A. (2004). Improving sustainability of buildings through a performance-based design approach. *2004 Preprint, National Renewable Energy Laboratory (NREL)* NREL/CP-550-36276.
- Dorst, K. (2008) Design Research: A Revolution-waiting-to-happen, *Design Studies*, vol. 29, no. 1, pp 4-11.
- Dorst, K., & Cross, N. (2001). Creativity in the design process: Co-evolution of problemsolution. *Design Studies, 22*, 425-437.
- Earl, C., Johnson, J., & Eckert, C. (2005). Complexity, Chapter 7 *Design Process Improvement: A Review of Current Practice*.
- Eckert, C., Clarkson, J. eds. (2005). *Design Process Improvement: A Review of Current Practice*, Springer.
- Edvardsson, E., & Hansson, S. O. (2005). 'When is a goal rational?' *Social Choice and Welfare* 24, 343-361.
- Federal Energy Management Program (FEMP), (2007). Energy Independence And Security Act (EISA) of 2007 (pp. P.L. 110-140 (H.R.116.)).
- Gane, V., Haymaker, J., (2007). Conceptual Design of High-rises with Parameteric Methods. Predicting the Future, 25th eCAADe Conference Proceedings, ISBN 978-0-9541183-6-5 Frankfurt, Germany, pp 293-301.
- Gero, J. S. (1990). Design prototypes: A knowledge representation schema for design. *AI Magazine, Special issue on AI based design systems, 11*(4), 26-36.
- Gero, J. S. (1996). Creativity, emergence and evolution in design. *Knowledge-Based Systems, 9*, 435-448.
- Hudson, R. (2009). Parametric development of problem descriptions. *International Journal of Architectural Computing, 7*(2), 199-216.
- Ïpek, E., McKee, S., Caruana, R., de Supinski, B., & Schulz, M. (2006, October 21-25, 2006). *Efficiently exploring architectural design spaces via predictive modeling.* Paper presented at the Proceedings of the 12th international conference on Architectural support for programming languages and operating systems.
- ISO 10303-11:2004 Industrial automation systems and integration -- Product data representation and exchange -- Part 11: Description methods: The EXPRESS language reference manual.
- Kleijnen J., (1997). Sensitivity analysis and related analyses: a review of some statistical techniques. *J Stat Comput Simul* 1997;57(1–4): 111–42.
- Kotonya, G., & Sommerville, I. (1997). Requirements engineering: processes and techniques. Wiley, Chichester.
- Kunz, J., Maile, T. & Bazjanac, V. (2009). Summary of the Energy Analysis of the First year of the Stanford Jerry Yang & Akiko Yamazaki Environment & Energy (Y2E2) Building. *CIFE Technical Report #TR183.*
- Lamsweerde, A. (2001). Goal-oriented requirements engineering: A guided tour. *Proceedings RE'01, 5th IEEE International Symposium on Requirements Engineering*, 249-263.
- Lawrence Berkeley National Laboratory (LBNL) (1982). DOE-2 Engineers Manual, Version 2.1A. National Technical Information Service, Springfield Virginia, United States.
- Lawrence Berkeley National Laboratory (LBNL) (2008). EnergyPlus Engineering Reference, The Reference to EnergyPlus Calculations. 20 April, Berkeley, California, United States.
- Love, T. (2002). Constructing a coherent cross-disciplinary body of theory about designing and designs: some philosophical issues *Design Studies, 23*(3), 345-361.
- Maher, M. L., Poon, J., & Boulanger, S. (1996). Formalizing design exploration as co-evolution: A combined gene approach. *Advances in Formal Design Methods for CAD*.
- McManus, H., Richards, Ross, M., & Hastings, D. (2007). A Framework for incorporating "ilities" in tradespace studies. *AIAA Space.*
- Ostergaard, K., & Summers, J.(2009) Development of a systematic classification and taxonomy of collaborative design activities, *Journal of Engineering Design*, 20: 1, 57-81.
- Papamichael, LaPorta, & Chauvert. (1998). Building design advisor: automated integration of multiple simulation tools.*, 6*(4), 341-352.
- Papamichael, & Protzen. (1993). The limits of intelligence in design. *Proceedings of the Focus Symposium on Computer-Assisted Building Design Systems 4th International Symposium on Systems Research, Informatics and Cybernetics*.
- Payne, J., Bettman, J., & Schkade, D. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty, 19*, 1-3.
- Phoenix Integration. (2004). Design exploration and optimization solutions: Tools for exploring, analyzing, and optimizing engineering designs, Technical White Paper, Blacksburg, VA.
- Phadke, M. S., & Taguchi, G. (1987). Selection quality characteristics and s/n ratios for robust design. *Ohmsha Ltd*, 1002-1007.
- Ross, A. (2003). Multi-attribute tradespace exploration with concurrent design as a value-centric framework for space system architecture and design. *Dual-SM*.
- Ross, A. M., & Hastings, D. E. (2005). The tradespace exploration paradigm. *INCOSE 2005 International Symposium*.
- Ross, A. M., Hastings, D. E., Warmkessel, J. M., & Diller, N. P. (2004). Multi-attribute tradespace exploration as front end for effective space system design. *Journal of Spacecraft and Rockets, 41*(1).
- Shah, J., Vargas-Hernandez, N., & Smith, S. (2003). Metrics for measuring ideation effectiveness. *Design Studies, 24*(2), 111-134.
- Simpson, T., Lautenschlager, U., & Mistree, F. (1998). Mass customization in the age of information: The case for open engineering systems. *The information revolution: Current and future consequences*, 49-71.
- Simpson, T., Rosen, D., Allen, J. K., & Mistree, F. (1996). Metrics for assessing design freedom and information certainty in the early stages of design. *Proceeding of the 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference*.
- Smith, R., & Eppinger, R. (1997). A predictive model of sequential iteration in engineering design. *Management Science, 43*(8).
- Takeda, H., Veerkamp, P., Tomiyama, T., & Yoshikawa, H., (1990). Modeling Design Processes, *AI Magazine* Vol. 11, No. 4. pp. 37-48.
- Tribelsky, E., & Sacks, R., (2010). "Measuring information flow in the detailed design of construction projects." Res Eng Design, 21, 189-206.
- Tate, D., & Nordlund, M. (1998). A design process roadmap as a general tool for structuring and supporting design activities. *Journal of Integrated Design and Process, 2*(3), 11-19.
- Wang, W. C., & Liu, J. J. (2006). Modeling of design iterations through simulation. *Automation in Construction 15*(5), 589-603.
- Watson, I., & Perera, S. (1997). Case-based design: A review and analysis of building design applications. *Journal of Artificial Intelligence for engineering Design, Analysis and Manufacturing AIEDAM, 11*(1), 59-87.
- Welle., B & Haymaker, J. (2010) [Reference-Based Optimization Method \(RBOM\): Enabling](https://files.me.com/haymaker/pty0xu) Flexible Problem Formulation for [Product Model-Based Multidisciplinary Design](https://files.me.com/haymaker/pty0xu) [Optimization.](https://files.me.com/haymaker/pty0xu) Available from: http://cife.stanford.edu/online.publications/TR197.pdf [Accessed January, 2011].
- Woodbury, R. F. & Burrow, A. L. (2006)*. "Whither design science?" Artificial Intelligence for Engineering Design Analysis and Manufacturing 20(2): 63-82.*

Appendix

Variables	Options		
		Additional	
	Final Schematic Design	alternatives	ASHRAE Baseline
Building Geometry	A	B, C	A
Windows	U-value: 0.30; SC: 0.44;		U-value: 0.57; SC: 0.57;
	SHCG: 0.378		SHCG: 0.49
Roof	U-value: 0.33		U-value: 0.65
Wall, Above Grade	U-value: 0.083 (above grade)		U-value: 0.113
Wall Below Grade	U-value: 0.056		U-value: 0.1
Percent Glass on	95%		40%
Exterior			
Percent Exterior with	50%	38%	0%
Shading			
Electric Lighting	1.10 w/sf		1.22 w/sf
Daylight Sensors	Yes		N ₀
HVAC	B:	A, C	B:
	High efficiency, Heat recovery		Standard Efficiency,
	Outside air minimum: 10%		Outside air minimum:
			20%

Table 3: Options for Variables explored in professional Design Process

For the purposes of comparison in our case studies, above and below grade wall Variables are modeled together.

Table 4: Options for Variables explored in advanced Design Process

Variables	Options		
Building Geometry	Square $(100 \text{ft} \times 100 \text{ft})$	Rectangular (200ft x 50ft)	
Building Orientation	$-90, -45, 0, 45, 90$		
Window Construction	U-value: 0.30	U-value: 0.57	
	SC: 0.44	SC: 0.57	
	SHCG: 0.378	SHCG: 0.49	
Exterior Wall	U-value: 0.083	U-value: 0.113	
Percent Glass on	95%	40%	
Exterior			
Percent Exterior Shading	50%	0%	
Electric Lighting	1.10 w/sf	1.22 w/sf	
Daylight Sensors	Yes	N ₀	
HVAC	High efficiency, Heat	Standard Efficiency,	
	recovery, Outside air	Outside air minimum: 20%	
	minimum: 10%		

*alternative*s = 2*5*2*2*2*2*2*2*2 = 1280

Figure 8: Asymmetries used to determine ASI for Advanced Design Process. First order interaction between all combinations of Variables. Graphs generated by PIDO technology.

Figure 9: Estimated Impact on annual operating costs and first cost by Variable. Diagrams, generated by PIDO, show that one Variable (window construction) out of nine has competing Impacts on project Goals (energy use vs. first cost). These results are used to determine ISC.

Figure 10: Estimated percentage impact on Net Present Value ranked by Variable. Graph generated by PIDO. The following series of "importance percentages" were used to calculate VSD 0,1,4,8,10,14,15,16,32 for our case study according to the following equation.

.68 48 32 $100/9$ $\frac{11.2 - 10}{4} \times \frac{32}{4} =$