Developing a Measure of Quality for Engineering Design Artifacts

Molly Hathaway Goldstein[1], Camilo Vieira Mejia[2], Robin Adams1, Şenay Purzer[1], Mitch Zielinski[1]
[1]School of Engineering Education, Computer and Information Technology[2]
Purdue University,
West Lafayette, IN
goldstm@purdue.edu, cvieira@purdue.edu spurzer@purdue.edu,

***Abstract*—Design is recognized as a key engineering activity, and engineering is a fundamental part of science education at the K-12 level. However, it is difficult to assess student designs when the range of "correct" answers is wide. Feedback in the form of assessment helps students learn from a design activity and can direct students along the pathway of improvement. The purpose of this paper is to develop an assessment protocol to measure solution quality taking into account both objective and subjective design criteria (e.g. measurements of cost/energy used along with aesthetics). Three protocols are developed by analyzing and comparing 109 high school students' design solutions to a zero-energy home design task. Lessons learned from the first two approaches informed the third approach that highlights the importance of balancing trade-offs. Results suggest the Trade-off Value approach provides an intuitive and accurate way to understand how well a designer has balanced both complementary and competing design criteria. These results can be used as feedback to support a systems design process and to evaluate the relationship between design quality and design behavior.**

Keywords— engineering design, design assessment, quality, trade-off

I. Introduction – Engineering Design & Quality

It is difficult to assess student designs where the range of "correct" answers is wide. Within an engineering context, design is "a systematic, intelligent process in which designers generate, evaluate, and specify concepts for devices, systems, or processes whose form and function achieve clients' objectives or users' needs while satisfying a specified set of constraints" [1, p104]. Here, design involves attending to an interconnected system of non-negotiable (e.g., natural laws) and negotiable (e.g., social, political, economic, aesthetic, etc.) issues [2]. In this way, design is different from well-structured problem solving because design does not fully fit within a rational problem solving paradigm that emphasizes decomposition and optimization, but rather a situative and reflective practice paradigm that emphasizes finding problem-solution pairs that *balance* conflicting objectives, needs and constraints [3].

A need to better understand design quality has been recognized globally from both the design community and from stakeholders in design projects [4] stemming from the concern that focusing on reducing cost and time on projects could lead to a "loss of functionality and boring, unattractive design" [4, p319]. However, the authors note that measuring overall design quality presents major conceptual and practical problems.

This paper addresses the need for a protocol for measuring design quality through this balancing trade-offs lens. Through the context of a design project directed towards a net-zero energy home, design quality is investigated through three methods. To begin, the design challenge is introduced. Next, two mathematical approaches to determining quality are shown with results, and are compared to two experts' evaluation. Lessons learned from these two methods informed a third approach – the Trade-off Value method. We then describe both the conceptual model and preliminary results, and identify future work.

II. Design Challenge

A. Participants

The study participants were 109 high school students enrolled in a 9th grade (ages 14-15) physical science course. Data were collected in five sections of the class. The majority of students were high school freshman (99.1%), with a nearly even number of females as males (49.5% female, 46.8% male, and 3.7% preferring not to answer). Of the 109 students, design files (logged process data) were complete for 84 students.

Students completed the Energy-Plus Home Design (http://energy.concord.org/energy3d), which prompted students to design an eco-friendly home by coming up with three unique designs that would consume minimal energy over the course of a year. The system offers a variety of simulation tools to analyze energy consumption and construction cost. The house is to cost less than $60,000, and has specified dimensions depending on house style type, as shown in Figure 1.

Colonial	Cape Cod	Ranch
Budget <$60,000, area 120-160 m^2, height 8-10 m	Budget <$60,000, area 100-150 m^2, height 7-9 m	Budget <$60,000, area 150-200 m^2, height 4-6 m

Fig. 1. Energy-Plus Home specifed house types

In addition to cost and size specifications, students were given the following eight constraints:

TABLE I. ENERGY-PLUS HOME DESIGN CONSTRAINTS

Each side of the house must have at least one window.
Tree trunks must be at two meters away from the walls of the house.
Do NOT add more than 40 solar panels.
Roof overhang must be less than 50 centimeters wide (the default is 25 centimeters).
Doors cannot be wider than two meters
Doors cannot be taller than three meters.
Keep the room temperature of the house to be 20°C all the time.
Keep height for house within specified range for house type.

B. Data Sources

As part of the design challenge, students needed to create three different types of house designs: (1) Ranch; (2) Cape Cod; (3) Colonial. Students were free to choose when to start working on one design or another one. Using the process data generated from Energy3D, the research team determined that students would spend the majority of their time (i.e. more than 100 minutes out of approximately 200 minutes) on one house design, while spending significantly less time on the other two designs. Eight students spent most of their time creating a Ranch, 34 students creating a Cape Cod, and 42 students a Colonial house. The following data was collected from theses final designs: (1) total annual energy consumption, (2) total construction costs, (3) approximate volume of home, (4) total area of all windows of home, (5) total area of all walls of home, and (6) total number of satisfied constraints out of 8.

III. "OVERALL" QUALITY SCORES

Two approaches were evaluated as an overall quality assessment of the design task. The first approach compared the *net-values* given by the simulations of the CAD tool for each student to the extreme performances in the sample overall (e.g. maximum energy consumption, maximum cost). The second approach ranked student's design using *performance percentiles* to compare them to each other, and integrate the different criteria within the sample scale. For reference, the designs in our sample had construction costs ranging from $48,523 to $66,610 and annual energy consumption values ranging from -7,861 kWh to 5,778 kWh. Both approaches draw on the same data, in an effort to leverage the "big data" affordances of data generated with in the CAD system, but take different approaches to producing an overall quality score. In particular, the net-value approach involves weighting the relationship between components and the percentile approach involves a mathematical normalization process within a given sample.

A. Net value approach

The net value approach involved an equation that would take into account the net performance values from simulations offered by the CAD tool. Given the context of the design challenge, we included the following variables: annual net energy consumption, cost of house materials, livability, and number of unsatisfied constraints. Equation 1 represents the general model of this approach.

$$Q_x = k_1 * Energy_x + k_2 * Cost_x + k_3 * Livability_x - k_4 * USC_x$$

Where,

$$Energy_x = \frac{V_x}{max(V)} * \left(\frac{\sum_i WindowArea_i}{\sum_j WallArea_j} * \left(l - \frac{E_x}{max(E)} \right) \right)$$

(1)

$$Cost_x = \frac{V_x}{max(V)} * \left(l - \frac{C_x}{max(C)} \right)$$

$$Livability_x = \left(\frac{\sum_i WindowArea_i}{\sum_j WallArea_j} + \frac{NTrees}{max(NTrees)} \right)$$

$$USC_x = \frac{\#\ of\ unsatisfied\ constraints}{\#\ of\ constraints}$$

Note that k_i represents a weighting coefficient for each of the elements in the equation. The energy, cost, and livability scores are computed relative to the maximum net values from students' designs (e.g. NTrees/max(NTrees)).

The energy efficiency of the house depends directly on the size of the house (i.e., volume V_x), and the area of the walls that is covered by windows (*WindowArea*). To control for these dependencies, the energy score is multiplied by the relative volume of the house ($V_x/max(v)$), and by the window-wall ratio.

B. Percentile Approach

The second approach considered the same four elements from the first equation but instead of using the net values in the equation, we first compared them to the rest of the sample to identify the percentile rank to which they belonged. This percentile (i.e a value from 0-1) was multiplied by 10 so that all four components would have the same scale (i.e., 1-10). Hence, the integration of these components was made by multiplying them inside a fourth root, as depicted in equation 2.

$$Q + x = (Energyu_x * Cost_x * Livability_x * USC_x) \quad (2)$$

C. Calibration

In order to calibrate our quality approaches, we asked two experts to rank students' designs according to their experience and preferences. The experts were provided with single-page summaries for each design that included two pictures of the designed house, and a table with the summary of the performance measures: annual energy consumption, cost, area, and volume of the house, number of trees, window area to wall area ratio, and whether the design satisfied each of the constraints.

The two experts organized the student designs in different piles based on what they considered as high or low quality design. We created two sets of weighting coefficients for the net-value approach based on these piles, trying to get as close as possible to experts' criteria.

Note that the percentile approach assumes that all components have the same weight. The *net-value approach* uses different ranges of values for different sections of Equation 1. Table II depicts the resulting values employed as weighting coefficients.

TABLE II. WEIGHTING COEFFICIENTS

Table Head	Table Column Head			
	k1 - Energy	***k2 - Cost***	***k3 - Aesthetics***	***k4 - Constraints***
Net-value_1	1	1	1	0.5
Net-value_2	1	1.5	0.5	0.5

D. Comparing Net-Value and Percentile Approaches

After computing the quality score using different weighting coefficients, we organized designs based on a 0-1 scale, where 1 corresponded to the best design and 0 corresponded to the worst one. This latter step was necessary in order to compare different approaches under the same scale.

1) Ranch

Eight students (C03, C12, D15, D21, E07, E13, G12, and G22) spent most of their time working on the Ranch design. Figure 2 shows how each ranch design differed in quality score among the three approaches. While the general trend appears comparable, there are notable differences. The highest and lowest scores for the *net-value* approach were D15 and E07 respectively, while the highest and lowest scores for the *percentile* approach were G12 and D21. In the net-value approach, the different weighting coefficients did not have an impact for the lowest and highest scores. For example, student C03 performed relatively well regardless of the quality assessment approach while D15 varied substantially based on quality score calculation method.

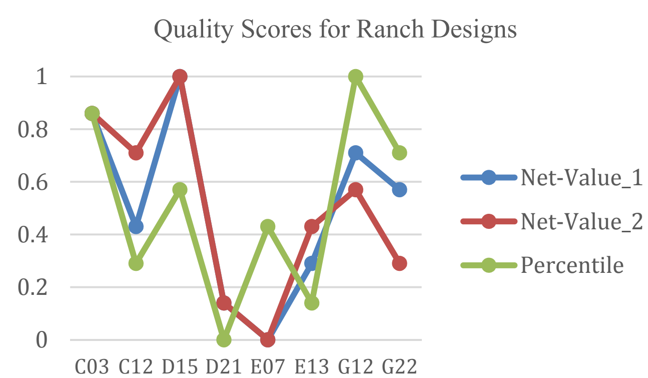

Fig. 2. Quality Score for Ranch Designs

2) Colonial

Forty-two students spent most of their time working on the Colonial design. We divided the scoring results in two different line plots (see Figures 3 and 4) to better show the differences among the different approaches. The highest and lowest scores for the *net-value* approach were students D17 and D20 correspondingly. D17 was below the average score for the *percentile* approach, and D20 was above the 20%. The highest and lowest scores for the *percentile* approach were students F12 and E14 respectively. Again, the *net-value* approach did not show differences between for the lowest and highest scores using different weighting coefficients, but it did show difference for students like C06 or E09.

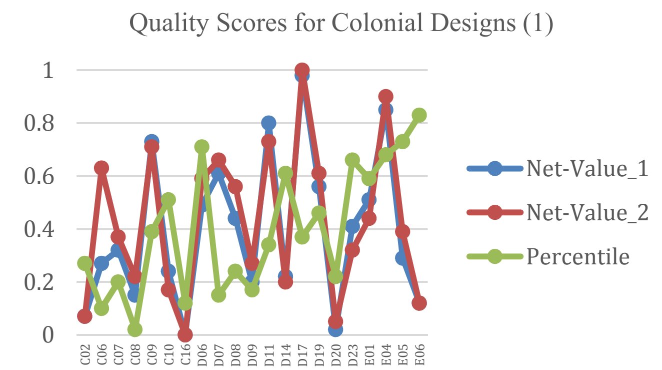

Fig. 3. Quality Score for Colonial Designs (first half class)

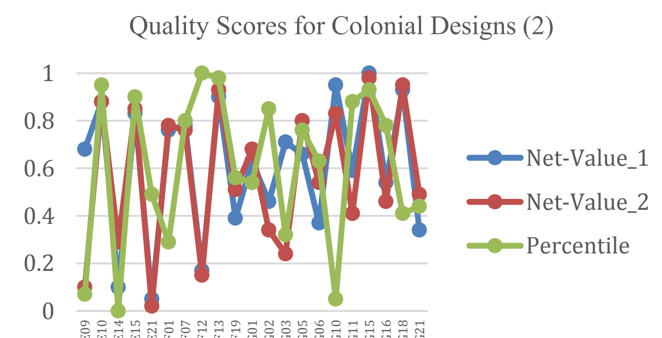

Fig. 4. Quality Score for Colonial Designs (second half of class)

Student D17's design is show in Figure 5. While this design scored high using the *net-value* approach it faired below average for the *percentile* approach. Construction costs for the design is $58,191 and annual energy consumption is 2,665kWh. The cost is below the threshold, but the energy consumption did not meet the goal of zero-energy as did many other designs. However, as seen from Figure 5, the house met attributes of aesthetics and livability with the large number and placement of reasonably sized windows.

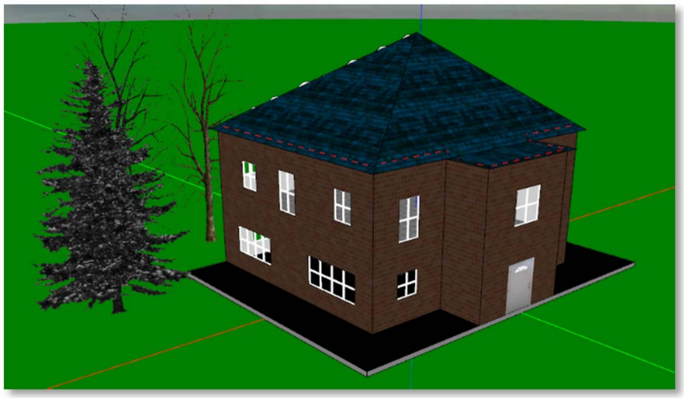

Fig. 5. Student D17 Colonial Design

3) Cape Cod

Thirty-four students focused on the Cape Cod design. Unlike the previous scenarios, the highest and lowest score for the net-value approach varied for the two set of weighting coefficients. Net-value_1 showed the highest score for G07 and the lowest score for D05, while for Net-Value_2 the students G04 and D04 occupied those positions. The *percentile* approach identified C05 as the highest student design in the sample, and D16 as the lowest quality design.

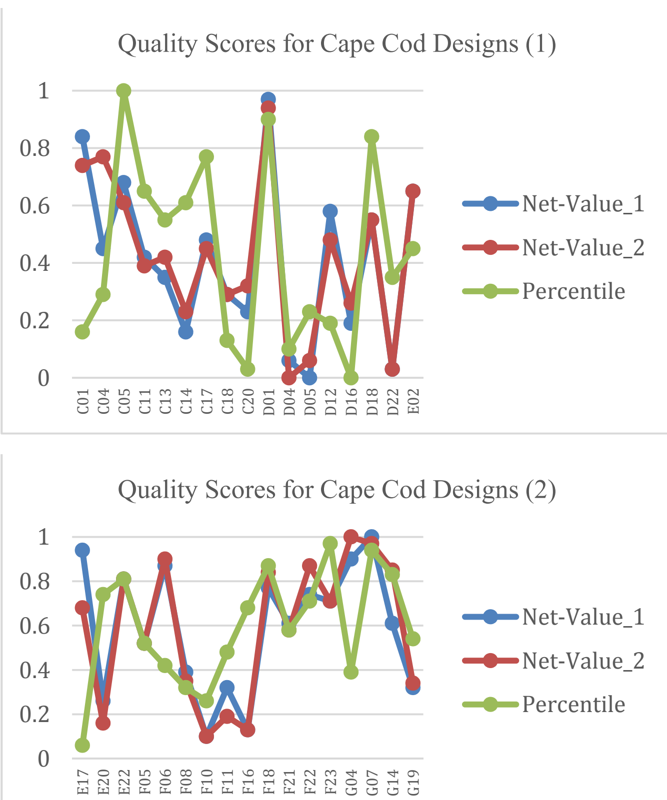

Fig. 6. Quality Score for Cape Cod Designs

E. Lessons Learned from Computing an Overall Quality Score

Two different approaches were evaluated with a group of 84 students working on different designs. The *net-value* approach was calibrated based on experts' quality criteria with two different sets of weighting coefficients aiming at giving higher or lower relevance to certain factors.

The results show that different approaches and different weighting coefficients within the same approach would lead to different or even contradictory results. Although the weighting coefficients were "*calibrated*", this calibration is context-dependent, and involves a certain degree of subjectivity. In addition, results across the three house styles indicate that the ranch and colonial quality scores appear to be different in nature than the Cape Cod quality scores.

These studies indicate that taking an "overall" quality approach to evaluating a design solution did not resolve our early concerns regarding repeatability, subjectivity, and credibility. Even though we took a very systematic approach, the open-endedness of the tasks requires students to continually make trade-off decisions to find a balanced solution that meets the criteria. Each of the quality assessment approaches and weighting coefficients benefit one of the factors more than the other ones, and therefore, the results were found to be inconclusive.

Therefore, we revisited the expert review process used to generate weighting coefficients. In reflecting on their process, the expert judges noted they had some variations in their ratings dependent on what they found important. Rather than labeling this as "subjectivity", we noticed that the considerations driving many of these ratings involved balancing trade-offs in the designs, even if this wasn't explicit. For example, they tried to understand the relationship in overall quality between energy consumption and total cost. In some cases this meant creating sub-piles of designs that did not meet one or more criteria (e.g., low energy consumption or total cost), but met other criteria including the more subjective criteria. We also noticed that the biggest factor contributing to livability to the judges was the presence of windows that would allow sunlight to enter the home; designs that had a "dungeon" quality with little sunlight received lower ratings. Similarly, designs that experimented with the overall shape of the house (beyond a four wall square or rectangle) were likely to receive "bonus points". In many cases, these "livability" issues provided a first step towards ranking designs, followed by an iterative approach of seeing the whole picture of how a design met (or didn't meet) all the criteria. Additionally, the judges confirmed that the presence of trees did not impact final ratings, including assessments of overall aesthetics. It should be noted that most, if not all, designs included trees, and as such the details of tree placement did not substantively distinguish designs.

IV. TRADE-OFF VALUE QUALITY SCORES

Understanding the extent to which a designer is able to balance conflicting objectives in a design provides one way to discuss quality of design. Balancing benefits and trade-offs is a behavior that experts designers engage in, and that students are less likely to exhibit as novice designers [5]. People in general, not only students, struggle with evaluating trade-offs between two attributes of very unequal importance where the option should be "obvious" [6] demonstrating that decision-making is not necessarily a naturally easy process.

A. *Conceptual Framework*

Engineering design includes a "synthesis of technical, human, and economic factors; and it requires the consideration of social, political, and other factors whenever they are relevant" [7, p2]. Here, human refers to more than ergonomics by encompassing what humans want, technical refers the design performance often through science and math concepts, and economic refers to monetary costs. While all of these areas of design are important, the level of emphasis on each can be quite subjective. Consider two different examples of coffee machines. In the first machine, say a Mr. Coffee®, cost (economic factors) is weighted much more heavily than a technical or human factors, such as heating elements and their relationship to flavor (see Figure 7). In the second example, a Technivorm Moccamaster®, human factors (the sleek appearance) in combination with technical factors (9-hole spray arm for an even soak of beans prior to extraction) are clearly weighted more than the cost (See Figure 8).

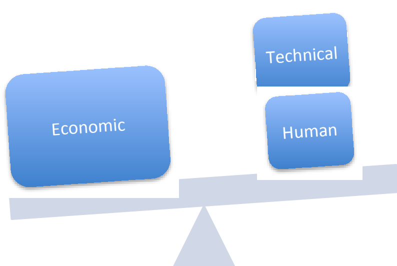

Fig. 7. Mr. Coffee® design trade-offs

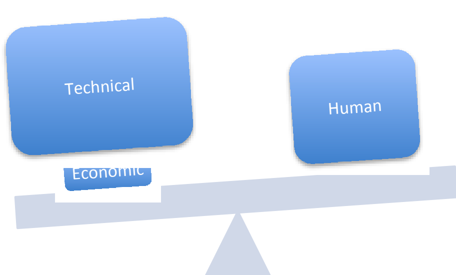

Fig. 8. Moccamaster® design trade-offs

The tradeoff value presented in Figure 9 is based on Asimow's [7] characterization of the distinct factors in engineering design, and is defined as the extent to which a design balances these competing factors. A design with a high trade-off value is able to take a systems approach to design, allowing consideration paid to the competing factors rather than focusing soley on optimizing one or two of the factors. Figure 9 demonstrates the overlap of the factors as the trade-off value. This idea complements more current views such as experience design by (balancing trade-offs among desirability, feasibility, and viability) [8] (that emphasize a human-centered perspective.

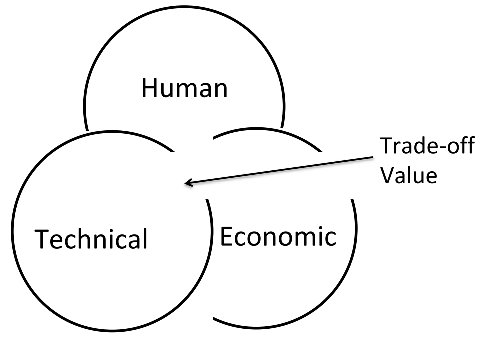

Fig. 9. Trade-off value conceptual framework

B. *Trade-Off Value Approach*

To understand the trade-offs students were making we calculated a value for each human, technical, and economic feature of each student's design. Table III summarizes how the log data were mapped to the conceptual framework. As shown here, some calculations (e.g., volume of home) map to multiple criteria, demonstrating system inter-relationships among size of the house, energy consumption, and total construction cost.

TABLE III. TRADE-OFF FRAMWORK WITH DATA

Engineering Factors	**Representative Design Data**
Technical	Total annual energy consumption (energy/volume) Approximate volume of home Number of satisfied constraints (out of 8)
Human	Total area of all windows of home (e.g., "livability" as demonstrated through Window-to-wall ratio for house)
Economic	Total construction costs (cost/volume) Approximate volume of home

We calculated the energy/volume and cost/volume for each design to standardize the energy and cost figures to account for distinct house styles/sizes. We converted each of the four design data from Table III (i.e., energy/volume, number of met constraints, window-wall ratio, and cost/volume) to percentile rank, ensuring that the rank order was logical in Equation 3. For example, a high construction cost/volume would be undesirable. Each percentile rank corresponds with a number out of 100.

$$Trade-off\_Value = \sum_i^j rank(human\_factors) + \sum_i^j rank(technical\_factors) + \sum_i^j rank(economic \quad (3)$$

$$Where$$

$$Human\ Factors = Livability = \frac{\sum_i WindowArea_i}{\sum_j WallArea_j}$$

$$Technical\ Factors_1 = \frac{Energy_i}{vol_i}$$

$$Technical\ Factors_2 = \frac{\#\ of\ unsatisfied\ constraints}{\#\ of\ constraints}$$

$$Economic\ Factors = \frac{Construction\ Cost_i}{vol_i}$$

C. *Trade-off Value Preliminary Findings*

Trade-off values in the sample ranged from 15.92 to 65.92 with a mean of 50.60 and median of 15.63 out of a possible score of 100. Table IV summarizes the factors that contributed to the trade-off value for the Highest and Lowest quality scores for the sample. Images of these homes are also included and discussed. Overall, the higher Trade-off value appears to indicate the extent to which all constraints were satisfied, including competing constraints. For example, increasing the window-to-wall ratio will have a positive impact on overall cost and energy consumption but a negative impact on livability. Similarly, the lowest Trade-off value appears to indicate an emphasis on optimizing energy consumption and cost at the expense of other criteria.

TABLE IV. HIGHEST VS. LOWEST RANKED DESIGNS

	F13 (Highest)	E14 (Lowest)
Trade-off Value	65.93	15.92
Annual energy consumption (kWh)	-1,099	-17
Area (m^2)	121.1	127.9
Cost/Volume ($\$/M^3$)	88.74	114.75
Window-to-Wall Ratio	0.08	0.009
# Satisfied constraints (out of 8)	8	7

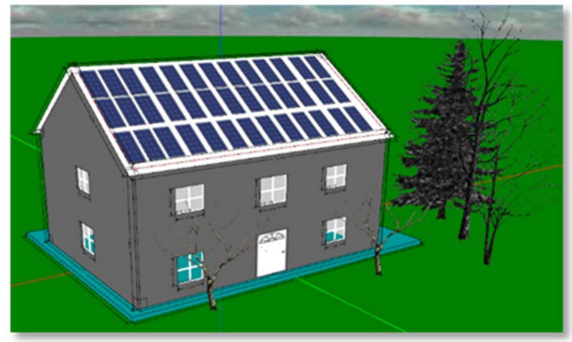

Fig. 10. Student F13 (Highest trade-off value) design view 1

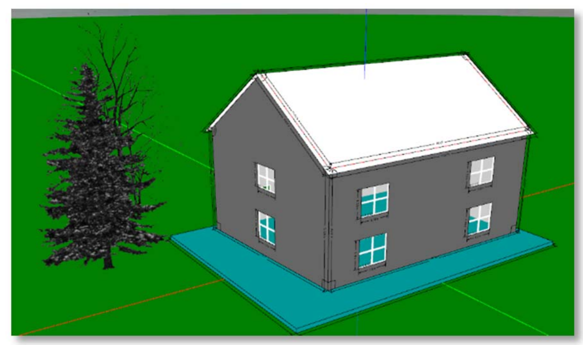

Fig. 11. Student F13 (Highest trade-off value) design view 1

While Design F13 had a low annual energy consumption (-1,099 kWh), the lowest in the class was -7,861 kWh. Additionally, the cost of F13 was competitive at $59,452 but the lowest in the class was $48,523. Conversely, E14 had final design criteria of -17 kWh and $53,604 compared to the highest and lowest energy consumption of 5,778 kWh and -7,861 kWh, and highest cost of $66,610. Accounting for the additional factors of livability (as demonstrated from window-to-wall ratio) and the degree to which constraints were met adds an additional level of complexity.

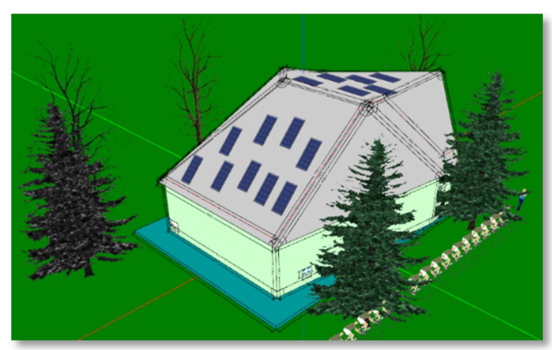

Fig. 12. Student E14 (Lowest trade-off value) design view 1

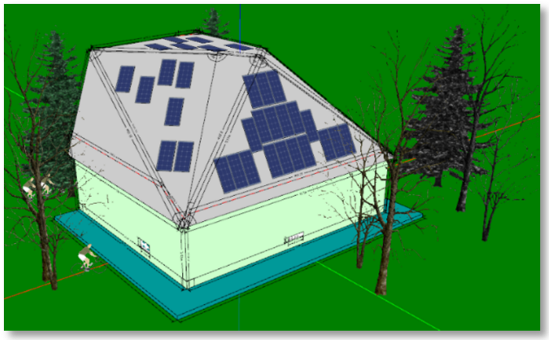

Fig. 13. Student E14 (Lowest trade-off value) design view 1. The small features on the bottom of each wall are windows.

D. Lessons Learned from Computing Trade-Off Value Quality Score

As a tool for thinking about design value, the *trade-off value* method of assessing design quality provides a comprehensive way to think about the interaction of client/user priorities, design possibilities and objective measures than attempting to assign an absolute measure of design quality. Trade-off value attempts to understand the extent of how well students were able to work within multiple conflicting areas of priority, taking a more systems approach to understanding design. Because this approach uses data that is stored in the log files, it would be allow for large-scale calculation of value of many designs. Initial results imply that variation of ratings would be reduced.

V. Conclusion & Future Work

Future work will compute the trade-off value for a larger number of student designs and will compared to the other overall quality approaches to determine if the trade-off value approach continues to better replicate the experts design review process.

VI. Limitations

This study is an attempt to broaden our understanding and assessment of design quality. While we have explained some of the limitations of assessing quality that ultimately led to us assessing trade-off value, there are limitations with the trade-off value method. It is difficult to assess aesthetics and livability on a large scale. While expert reviews would provide a more thorough appraisal of human factors, the window-to-wall ratio attempts to highlight the feature our experts found key for livability.

VII. Acknowledgement

This work presented in this manuscript is based upon work supported by the National Science Foundation (NSF) under Grant DUE #1348547 and DUE #1348530. Any opinions, findings, and conclusions or recommendations expressed in this paper, however, are those of the authors and do not necessarily reflect the views of the NSF.

References

[1] C.L. Dym et al., "Engineering Design Thinking, Teaching, and Learning", *J. Engineering Education*, vol. 94, no. 1, pp. 103-120, Jan. 2005.

[2] V. Goel and P. Pirolli, "The Structure of Design Problem Spaces." *Cogn Sci,* vol. 16, pp. 395-429. 1992

[3] K. Dorst, "Design Problems and Design Paradoxes." *Design Issues*, vol. 22, no. 3, pp. 4-17. 2006.

[4] D. Gann et al., "Design Quality Indicator as a tool for thinking", *Building Research & Information*, vol. 31, no. 5, pp. 318-333, 2003.

[5] D. P. Crismond and R. S. Adams, "The Informed Design Teaching and Learning Matrix", *J. Engineering Education*, vol. 101, no. 4, pp. 738-797, Oct. 2012.

[6] M. Scholten and S. J. Sherman, "Tradeoffs and theory: The double-mediation model", *J. Experimental Psychology: General*, vol. 135(2), pp. 237-261, May 2006.

[7] M. Asimow, *Introduction to Design*, Englewood Cliffs, NJ: Prentice-Hall, 1962.

[8] T. Brown, "Change by design," 2009.