2006-1452: EXPERIMENTAL VALIDATION OF AN IMPROVED DESIGN METHOD

Durward Sobek, Montana State University
Durward Sobek is an Associate Professor in Industrial and Management Engineering at Montana State University. He holds an AB degree in Engineering Science from Dartmouth College, and MS and PhD degrees in Industrial and Operations Engineering from the University of Michigan.

Joshua Ruder, Montana State University
Joshua Ruder is a Master's student in Industrial and Management Engineering at Montana State University.

Experimental Validation of an Improved Design Method

Abstract

Prior investigation found a statistical association between engineering effort aimed at system-level design issues and the quality of design outcomes in senior design projects, but that simply "telling" students to consider system architecture and interface issues in their designs was not effective. We developed a method to help design engineers with this important phase of design, then conducted an experiment involving mechanical engineering students to test its effectiveness. This paper describes the experimental method, presents results, and discusses the implications for engineering education research.

Introduction

In prior research on student design projects, we observed that system-level design effort associates positively with productivity and design quality. System-level design pertains to questions of product or system architecture, configuration, and layout; and as such, provides an important bridge between conceptual design work and detailed design decisions. While prior research indicates that this phase of design seems important to successful outcomes of design processes, it is not well understood and we have not yet established a causal link. To do this, we developed a tool designed to elicit system-level design work from the user. We then conducted an experiment to test whether use of the tool improves design performance among student designers.

Many methods and tools have been developed to teach good practices and assist the design process.[1] Existing tools and methods tend to focus on either the front-end tasks of conceptual design, particularly ideation and selection, or on the back-end, such as computer-aided design and analysis tools. However, little is available to guide the phase that bridges concept and detailed design, what we call system-level design. We hope to address this gap.

Interestingly, few of the existing tools and methods have been verified experimentally to validate the superiority of one method over another. It would seem that engineering educators would want to know whether a particular tool has been tested and verified as effective before presenting it to students as good design practice.

Thus, this paper contributes to engineering education on two fronts. First, we provide a method for improving student design processes that has been experimentally validated, which would be of interest to educators interested in engineering design. Second, we describe a cross-over experimental method which can be useful to a broad range of education researchers wanting to test pedagogical tools/methods experimentally. The experimental design has simple but strong internal and external validation indicators, and overcomes some of the ethical issues which often surround experiments in an educational setting.

Background

The first design phase following need identification is generally concept design; that is, addressing a given problem with preliminary ideas, strategies and approaches. This phase includes problem definition, information gathering, and idea generation activities. It usually concludes with the selection of the solution approach that will be pursued. Eventually, the design team will detail out the solution by quantifying the specific features needed to realize the concept. According to Pahl and Beitz [2], detail design "completes the embodiment of technical products with final instructions about the layout, forms, dimensions, and surface properties of all individual components."

An important set of design decisions lie between the conceptual and detailed design phases. We call this bridging set of activities system-level design (SLD), defined as: *exploration of and decisions about components and subsystems, and their configuration*. SLD starts with the solution approach decided in the conceptual design, and encompasses elements from embodiment design,[2] system architecture,[3,4] preliminary design,[5] product planning,[6] and modularity.[7] These decisions are extremely important to the overall success of a design project.

Interestingly, system level design has not been heavily studied. Some information is available from specific experiences of a designer or educator. These authors often state the importance of system level design, but do not supply a method or tool to fill that gap. For example, one design text states that this intermediate phase requires "a flexible approach with many iterations and changes of focus."[1] We feel that tools can be developed to help design teams balance the often complex and competing objectives involved in SLD.

Our research group identified a statistical correlation between SLD documented in student design journals and project outcomes from analyses that attempt to isolate design process elements that contribute to desirable design outcomes.[8-11] This research indicated a possible causal relationship that could be tested through experimentation. An initial screening experiment[12] strengthened the correlation in an experimental environment, and pointed toward the need for a system level design tool that focuses on the design of interfaces between functional subsystems. Based upon prior work, we felt that the application of the tool prior to committing to a final conceptual design would result in better design quality.

Proposed Design Method

Morphological analysis is a well-known technique for generating ideas.[13,14] The designer identifies the sub-functions needed to meet the stated design requirements, then brainstorms different ways the sub-functions can be accomplished. By making different combinations of the sub-function alternatives, the designer generates a large number of overall conceptual ideas.

We adapted the morphological approach to explicitly investigate system-level design considerations (see Figure 1). For each conceptual design idea under consideration, the design team identifies the key functions that the concept must execute to achieve the overall design objective. Multiple implementation options are then depicted for each of the functions in a matrix format. The options are likely to have unique interface configuration requirements when combined with the other functional options. Additionally, some options may be incompatible with others, while some may require a specific option of another function. Once these considerations have been fleshed out, the designer generates a list of potentially feasible alternative configurations for this particular concept. One way to generate alternative configurations is to create combinations of options that optimize each of the different functions.

Figure 1: Overview of Proposed Design Method

1. Generate conceptual design alternatives
  a. Define the problem
  b. Generate alternatives
  c. Narrow the set of alternatives to a manageable size

2. Apply System-Level Morphological Analysis to one alternative
  a. Identify the key functions of the concept needed to realize the design objectives
  b. Generate 2 or more options to accomplish each function
  c. Identify which function options cannot be used with other options (exclusions)
  d. Identify which function options require inclusion of another option (dependencies)
  e. Generate a list of alternative configurations for the concept alternative
  f. Investigate interface feasibilities
  g. Select the most promising configuration

3. Repeat step 2 for each alternative

4. Compare alternatives using the best configurations

5. Choose the best alternative

In order to select the best configuration for that conceptual design alternative, the design participants engage in a more lucid discussion of the interfaces, including those with the user, environment, or another device. The broader set of interfaces sets the stage for a story-telling walk-through of the design, describing the functional path of the design objectives in terms of the interface requirements. At each point the designer considers how the interfaces are handled and whether alternative methods might exist for meeting the interface requirements.

The above steps are repeated for each concept alternative under consideration. The result is a set of best configurations for each concept. The design team can then enter the concept selection decision with the best known configurations of the alternatives, rather than preconceived versions of the concepts. Inherently, then, the concept selection decision considers subsystem interface issues.

Experimental Design

Given prior research results, we hypothesized that:

> *Design processes which incorporate system-level morphological analysis will produce better designs than design processes which do not.*

Since the morphological tool and method were designed to elicit system-level design considerations, this hypothesis implies that design processes which systematically consider system-level issues will outperform processes which do not. The remainder of this section details an experiment to test the hypothesis among senior mechanical engineering students.

The participants were students enrolled in ME 403 Mechanical Engineering Design I. This course is structured as a design project experience emphasizing use of a formal design process, presentations, and documentation. The course also includes coverage of industry machining and welding practices.

The experiment was designed as a crossover design, as depicted in Figure 2. A crossover design is a special type of repeated measurement experiment where experimental units are given different treatments over time, and pretest data are compared to posttest data. In a crossover design each experimental unit serves as its own control. The comparisons of primary interest are the scores of the two problems between runs. Secondarily, we are interested in changes in performance between runs of the two groups, if comparability between problems can be established. The arrows in Figure 2 show the expected comparisons and the directions of the improvement we hypothesize. Note that no difference is expected between the golf ball and mouse problems in run 1 or run 2.

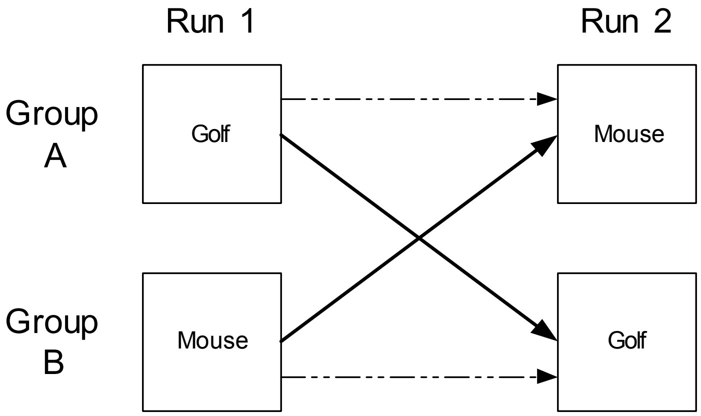

Figure 2: Graphical Depiction of Cross-Over Design

This type of design eliminates the ethical question of exposing a group of students to a potentially beneficial treatment without giving that same treatment to the control group. The experimental design also has good external validity. The comparisons allow for clear results that are either positive or negative, with little room for grey areas. This method allows for within group testing and randomization without the need for large sample sizes.

A disadvantage of this design, however, is that it is not a "true" experiment since we do not randomize the second run, making internal validity potentially less robust. In addition, bias may enter into the results due to participant learning between the experimental runs or other sequential effects. But these biases can be balanced by the timing and implementation of the design. In our case, we timed the experiment to coincide with classroom activities so as to minimize the effects of sequential learning from one run to the next, and of classroom learning that might occur between runs, as will be discussed later.

The design problems used in the experiment were crafted to be solvable within two hours, yet complex enough to have competing design objectives and not have obvious solutions. The first problem is that of moving a golf ball to a target while negotiating a drop of 25 cm. Design performance was measured by the final resting location of the golf ball on a target of concentric rings. The second problem was to transport a hacky-sack (the "mouse") to a linear target area that was strongly defined on one side and weakly on the other. In addition to performance measures, designs were also evaluated by the number of parts used, with fewer numbers of parts being better. The number of parts is a rough proxy for simplicity, a well-known axiom of design.

Both of these problems were solved using only the parts provided: an assortment of Lego parts, string, and a rubber band. The only difference between the problems was the amount of string (60 cm for the golf ball problem and 120 cm for the hacky-sack problem) and the number of wheels supplied (4 wheels for the golf ball problem and 6 wheels for the hacky-sack problem).

Experimental Protocol

During Run 1, the experimenter lead the design team through the following steps:

1. Introduction and familiarization exercise
2. Generation of at least three ideas
3. Selection of best idea
4. Prototype build and test
5. Final demonstration

Run 2 followed the same basic design process as Run 1, except for the following two changes:

- Familiarization exercise was eliminated.
- System-Level morphological analysis was introduced before concept selection.

Since the student participants were only superficially familiar with the new design tool and method (introduced in lecture the week before), the experimenter walked the participants through the sequence of steps and guided them in applying the tool to their problem. The experimenter was careful not to suggest design ideas or identify potential problems, but merely asked the participants to execute each step as indicated in the protocol.

The overall amount of time allocated to execute a run had been used in previous experiments. The time allocated allowed most teams to finish the design problem fairly comfortably within the timeframe; some groups finished a little early while some had to push to finish, but all groups completed a testable prototype.

Analysis and Results

The final analysis included data from 7 teams of two in each of groups A and B. The two measurable quantities from the experiment were the performance score of the team's prototype over three trial runs, and the number of Lego pieces used in the prototype. The two factors were normalized and averaged in order to obtain a single measure of design quality for each participant team ranging from 0 to 1, with 1 being the best possible combined score. Thus we considered higher scores on the combined measure as "better" than lower scores, an admittedly crude but nonetheless objective measure of design goodness.

When checking the normality assumption of the data using a normal probability plot, two potential outliers were identified. No assignable cause was found for the first data point. However the second data point traced to a group that resisted the morphological tool in Run 2, and rejected the results of the interface discussion resulting in the lowest score in the sample by a significant margin. This data point was removed as an outlier due to failure to follow protocol.

Tests for equal variance between the key categories of comparison found that, across all categories of comparison, variances tested as statistically equal according to two-sample F-tests for variances. Next, two-sample T-tests assuming equal means were used to test whether the run 2 results were higher than the run 1 results. Table 1 displays the means tests results.

Table 1: Results of T-test of Means Assuming Equal Variance

	Run 1	Run 2	Difference
Golf ball problem	0.584	0.853	0.269**
Mouse problem	0.433	0.740	0.307**
Within-run difference	0.152	0.113	
Group A	0.584	0.740	0.156
Group B	0.433	0.853	0.421**

* p-value ≤ 0.10, ** p-value ≤ 0.05

The first comparison determines whether any improvement occurred from one run to the next on the same problem. The golf ball problem saw a 46% increase in scores from run 1 to run 2 (p = 0.028), while the mouse transport problem experienced a 71% improvement between runs (p = 0.027). These results strongly support the stated hypothesis.

Next, we compared the run 1 scores against the run 2 scores of groups A and B. Group A showed a 26% improvement, but the difference was not strongly significant (p-value = 0.142). Group B, however, showed a 97% increase with a p-value of 0.004. The highly significant result supports the hypothesis that use of the SLD tool improves design performance while the non-significant result is not contraindicating.

The final comparison checked the comparability between problems. In the first run, the average difference in normalized scores between the golf and mouse problems was 0.152, with a p-value of 0.331. In the second run, the average difference of 0.113, with a p-value of 0.366. These tests indicate no statistical difference in difficulty between the two problems, although it appears that the mouse problem may be slightly less difficult.

Discussion

Overall, the variance and means tests conducted on the normalized scores of the experimental groups support the hypothesis of the experiment, while the non-significant result of the means test between runs for group A does not supply contradicting evidence against the hypothesis. Figure 2 displays these results graphically.

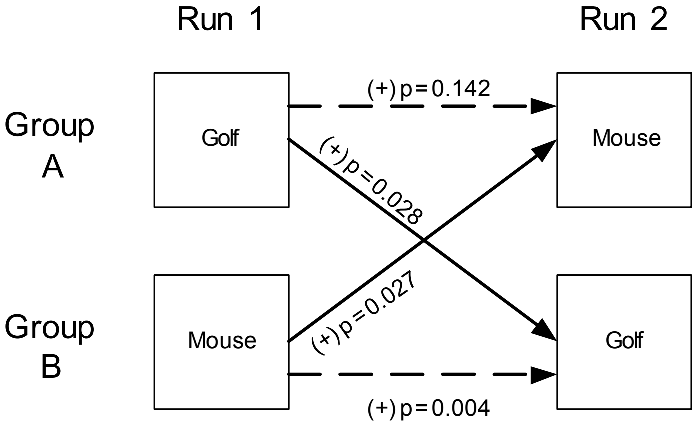

Figure 2: Graphical Summary of the Comparison Results

While the results seem to support the hypothesis that use of the system-level morphological analysis improves design performance, the experiment is not without potential biases. One concern is that participants could have learned something in the first run that helped them improve performance in the second. To counter this potential bias, we included a materials familiarization exercise at the beginning of the first run to minimize the effects of materials familiarity. We also separated the experimental runs by approximately five weeks (including spring break) to make it difficult to reconstruct the details of previous designs.

The hazard associated with this delay is the increased probability that students might learn something between runs that would aid their performance in the second run. While we cannot completely eliminate this possibility, we took several steps to minimize it. First, ME 403 classroom topics addressed items (such as catalog selection and machine shop tours) that would not likely apply to the experiment's design problems, save for training in the morphological tool used in the experiment and the theory behind it. Second, the experimenter carefully recorded observations and retained the participants' work products, and observed no indication that such learning occurred. For example, the students' level of ideation was comparable between runs, no comments referring to things learned in class or textbook were witnessed, and the only student comments referring to the first experiment were statements such as, "we can't do this the same way we did last time." Thirdly, we administered a post-experiment survey to the students before reporting the results: 89% of the participants felt that the design experience was worthwhile, 59% reported that the SLD tool was moderately to very helpful, and 59% would be willing to use the tool on a future project. While not conclusive, the survey results triangulate with the experimental results indicating that use of the SLD tool lead to improved performance.

A final source of possible bias is the interaction of the experimenter with the participants. In order to guide the students in use of the novel design tool, the experimenter necessarily interacted with the participants more in run 2 than in run 1. It is possible that simply interacting with an "expert" lead to superior results irrespective of the content of that interaction. However, the post-experiment survey results seem to indicate otherwise.

A topic for future work is the application of this method to problems of greater complexity. The problems used in this experiment were simple and straightforward. We hope that the tool's usefulness in discovering better designs will actually increase with the complexity of the design problems since interface issues become more critical with increasing complexity; but this remains to be tested. Another topic for future research is the applicability of the method to other design domains and populations. Is the method also useful for other areas of engineering design, such as electrical engineering? Is the method useful for extreme novice designers (e.g., first-year engineering students)? How about more experienced designers? To effectively address these questions requires a more fundamental understanding of why the tool is beneficial. The answers obtained would likely raise the further question of whether other SLD tools could or should be created to aid in the exploration and improvement of designs at a systems level.

Conclusions

The evidence presented here seems to provide fairly strong indication that system-level morphological analysis has a positive impact on the outcome of design. Improvement was seen across every comparison that was predicated in the hypothesis, and in 3 of the 4 cases, the improvement was statistically significant. However, the experiment was conducted on a fairly small population and more testing is needed.

The evidence suggests that use of this tool improves the outcome of the design process. We believe that the reason for the improvement is that the students using the tool came to a deeper understanding of the system level design issues that typically might not be addressed in the design selection process only to be discovered late in the process. From our observations of the participants, this deeper understanding made the selection of concept easier and less mysterious, which, in turn, aided in the transition between concept and detailed design, and decreased the amount of "tinkering" in the detailed design phase. In all, the students produced higher quality designs, as indicated by our performance and piece-count measures, by addressing the SLD issues raised in the system-level morphological analysis.

Perhaps just as importantly, this study provides a template for demonstrating the usefulness of an educational tool or method. Outlined in this paper is one approach for doing method comparisons that has a fairly straightforward protocol and simple diagnostics, yet contains statistically rigorous internal and external validation. This work will hopefully aid in the development and validation of additional tools and methods for use in engineering education.

Acknowledgements

Funding for this research was provided by the National Science Foundation, award # 9984484.

References

1. Wallace, K. and Burgess, S., Methods and Tools for Decision Making in Engineering Design. *Design Studies*, 1995, 16, 429-446.
2. Pahl, G. & Beitz, W., *Engineering Design: A Systematic Approach*, pp. 139 and 400, 1996 (Springer: London).
3. Sosa, M. E., Eppinger, S. D., & Rowles, C. M., Identifying Modular and Integrative Systems and Their Impact on Design Team Interactions. *Journal of Mechanical Design*, 2003, 125, 240-252.
4. Van Wie, M. J., Rajan, P., Campbell, M. I., Stone, R. B., & Wood, K. L., Representing Product Architecture. *ASME Design Engineering Technical Conference*, Chicago, Illinois, USA, 2004.
5. Dym, C. L. and Little, P., Engineering Design: A Project-Based Introduction, 2000 (John Wiley & Sons: New York).
6. Ulrich, K. T., and Eppinger, S. D., *Product Design and Development*, 2nd ed., 2000 (McGraw Hill: New York).
7. Gershenson, J. K., Prasad, Jagganath, G. and Zhang, Y., Product Modularity: Measures and Design Methods. *Journal of Engineering Design*, 2004, 15(1), 33-51.
8. Costa, R., & Sobek II, D. K., How Process Affects Performance. *ASME Design Engineering Technical Conference*, 2004.
9. Wilkening, S. H. and Sobek II, D. K., Relating Design Activity to Quality of Outcome. *ASME Design Engineering Technical Conference*, Salt Lake City, USA, 2004.
10. Sobek, II, D. K. & Jain, V. K., Process Factors Affecting Design Quality: A Virtual Design of Experiments Approach. Under review at *Journal of Mechanical Design.*
11. Jain, V. K. and Sobek, II, D. K., Linking Design Process to Customer Satisfaction Through Virtual Design of Experiments. Under review at *Research in Engineering Design*.
12. Ruder, J. A., & Sobek II, D. K., Student System Design Activities: An Empirical Pilot Study on Improving Design Outcomes," *ASME Design Engineering Technical Conference*, Long Beach, California, 2005.
13. Holt, J.E., The Designer's Judgment. *Design Studies*, 1997, 18, 113-123.
14. Marshall, R. and Leaney, P. G., Holonic Product Design: A Process for Modular Product Realization. *Journal of Engineering Design*, 2002, 13(4), 293-303.