Process Factors Affecting Design Quality:

A Virtual Design of Experiments Approach

Durward K. Sobek, II 1
Mechanical & Industrial Engineering Dept.
Montana State University
220 Roberts Hall
Bozeman, MT 59717-3800
USA
Tel: +1 406 994 7140
Fax: +1 406 994 6292
dsobek@ie.montana.edu

Vikas K. Jain
Isotec International, Inc.
201 Longview Road
Canton, GA 30169
USA
Tel: +1 800 234 6300, +1 404 232 9452
Fax: +1 770 479 1566
vikas.jain@isotecintl.com

January 6, 2005

1 Corresponding Author
ABSTRACT

This paper focuses on better understanding how design processes affect outcomes in mechanical engineering design. Following a general design research methodology, process data were collected from student design journals, which were then tested for association with design quality as measured by an external evaluation using a virtual design of experiments approach. The student teams that achieved higher quality designs placed greater emphasis on system-level design work and on concept-level problem definition activity, and lesser emphasis on concept-level idea generation. Whereas, detailed level design and design refinement activities associated with lower quality. These results confirm some recommendations and findings from the design literature, while also pointing to areas for further research.

Keywords: computational models, design education, design methodology
What is a “good” design process? Much work has been done and continues to be done to investigate this question. Some of the work is more prescriptive in nature, authors writing from their experiences about how design ought to proceed to achieve success. Increasingly, in last 10-15 years, design research has focused on describing design processes in actual and experimental settings to gain insight into what is a “good” design process.

However, as Blessing, Chakrabatri and Wallace [1] point out, for design research to continue to mature as a field, the growing body of research needs greater rigor, particularly in connecting the prescriptive and descriptive realms, and in connecting both to success measures. Blessing, et al., propose a general design research methodology where criteria for success are defined, descriptive studies understand design as it is currently practiced, prescriptive research defines new methods and tools, and additional descriptive studies validate the efficacy and appropriateness of new methods and tools. No one study is expected to cover all these steps, but every study can be framed in the context of the framework. Furthermore, Blessing, et al. [1] note that “systematic testing of methods and tools has not received much attention in design research, despite its importance” (p53).

The study reported here begins to address the gaps that Blessing and colleagues point out, in the mechanical engineering domain. We collected design process data from student engineering design projects, and carefully codified and quantified the data. In parallel, professional engineers independently evaluated the quality of the student project deliverables. We then statistically modeled the process data using the outcomes assessment as the response variable, explicitly connecting data descriptive of design
processes to measurable outcomes. In so doing, we attempt to validate certain findings from recent descriptive research and elements from the prescriptive literature. In several cases, the data did validate prior work; but the data also indicate some areas for additional research.

This paper presents the details of the approach and findings of this study. The next section provides a literature review as background. Then we describe the data collection and modeling methods used, contrasted against other methods commonly used to study design processes. Section 3 presents the modeling results, with discussion of those results and conclusions following in Sections 4 and 5.

1 BACKGROUND

Design has traditionally been an important part of an engineer’s training. The main accreditation organization for engineering programs in the United States (ABET) gives design significant emphasis in its evaluation criteria [2]. In fact, ABET requires that each accredited engineering curriculum have a capstone design experience. Design also plays an integral part in any organization with innovation as a core consideration. Thus, it comes as no surprise that in recent years, increased emphasis has been placed on design in engineering curricula. Even so, design may still be one of the least understood areas in engineering education.

Formal design research seems to have begun in the 1960’s, with so-called “first-generation” models used to attempt to find generic optimization routines that could be applied to any type of problem [3,4]. In 1969, Simon suggested that satisficing (i.e., working until the design is good enough) rather than optimizing might be a more appropriate approach [5], and over the next two decades this idea appears in the “second-
“third generation” models. During this time, two streams appear to develop in design research, with engineering researchers favoring heavily sequential design models, and architectural design researchers experimenting with more cyclical models. The architectural models also tended to include cognitive processes, while engineering models attempted to define stages in the design process. Third generation models arrived after the 1980’s, combining these two viewpoints [3]. Cross [6], Dym and Little [7], Haik [8], Pahl and Bietz [9], Pugh [10], and Ullman [11] represent a few examples of hybrid third generation models of engineering design.

This stream of work tends to be prescriptive. In each case, authors suggest that certain steps be followed if design is to be done properly so as to maximize the likelihood of a successful outcome, however that might be defined. The models tend to trade off precision in task definition with model stability with respect to sequence. Some of the earliest models (e.g., [12]) show very general steps like generate-conjecture-analyze, and simply say to repeat until done. Later models, like Ullman [11], have a detailed sequence prescribing the order in which a designer accomplishes everything from forming the design team to retiring the final product.

In our review of mechanical engineering design texts, similar to Blessing, et al.’s critique [1], we were unable to identify any design process models that had been empirically validated or that had explicitly correlated design process to outcome. Most authors seemed to be either expert designers writing from their work experience, or academics writing from their teaching experience. In either case, the proposed models do not claim to be based on rigorous research.
Interestingly, a number of recent studies indicate that actual design processes rarely follow the prescribed processes precisely. Maffin [13] observed from 12 case studies and numerous interviews conducted in industrial settings, that design activity in practice seldom resembles the engineering models advocated in the literature. He suggests that the models may not be useful in practice because they were developed in light of novel design problems (relatively rare in practice) and because they make unrealistic assumptions about the context of design (e.g., time constraints, availability of resources, quality of information). Other studies [14,15,16] observed that structured design methods seemed to be helpful if applied flexibly to allow for “opportunistic” behavior, the structured methods serving more as a general guide than a prescription.

Still, as Atman, et al. [17] aptly point out, the prevailing feeling is that how designers spend their time is as important as how much time they spend, perhaps more so. Much recent work in design research has aimed to gain insight into design processes through case studies and verbal protocol studies. Subjects range from first-year college students to highly experienced professional designers, and contexts include laboratory, educational, and industrial settings. This body of work has led to a number of insights relevant to our study of student engineering design processes.

One insight, as mentioned above, and which has quite a bit of support, is that structured approaches seem to help, but only if they are used flexibly to allow for opportunistic behavior. Ball, et al.’s study [14] of electrical engineering students found that a majority of transitions between activities were consistent with a top-down, depth-first structure, but that deviations from that structure seemed to occur at critical junctures. Pahl, et al. [15], in their summary of 12 years of design research in Germany, likewise
found that a “flexible-methodological procedure” was needed, as the successful designers they observed would adapt design processes to suit unique design problems. Bender and colleagues [16] found that design methodology education did not help engineering students until they had enough experience with the methodology to apply it flexibly. What’s less clear is what this means. Research on how the adapt these procedures (in what ways and when) is the subject of ongoing research.

A second insight with quite a bit of support is that problem scoping done well improves design success. Atman, et al. [17], and Adams, Turns, and Atman [18] in their protocol studies of first-year and senior-level engineering students consistently found that processes with high levels of information gathering, problem scoping, and problem setting activity tend to associate with higher design quality. Similarly, Pahl, et al. [15] and Badke-Schaub and Frankenberger [19] found that goal analysis is a key determinant of project success. Cross [20] notes that problem scoping seems to be associated with successful design projects across a number of expert designer studies.

On the other hand, the data on multiple competing alternatives seems to be mixed. Stauffer and Ullman [21] in their comparison of some early empirical studies in mechanical engineering design found conflicting results—some of the studies observed parallel solution search, others sequential solution search. Additional research does not seem to have brought much clarity, as some recent studies have associated solution search and generating alternatives with successful outcomes [15,17,19], while other studies observe that strategies involving multiple competing alternatives at any point in the design process are rare, and that satisficing behavior tends to dominate [13,14,20,22].
There appears to be little argument that domain expertise appears heavily influential [20,23], some claiming that it is more important than domain-independent procedural knowledge [21]. Fixation appears to be both common and undesirable [14,20]. And solution analysis appears associated with successful project outcomes [15,19].

Few of these studies, however, statistically validate observations against success criteria. Statistical analysis can be problematic given the small sample sizes characteristic of design research. Atman, et al. [17] use t-tests to identify statistically significant differences in activities between high scoring and low scoring designs, and use correlation statistics to identify statistically significant associations between activity variables and design scores. Bender, et al. [16] performed similar statistical analyses of the results from their quasi-experiment to look at differences between the control and experimental groups. These fairly simple analyses do attempt to draw explicit connections between design process and design outcome, but as we shall point out later, there is opportunity to use more powerful statistical modeling techniques to provide insight into the extent to which process variables can explain variance in project outcome.

An interesting vein of research has recently emerged on expert designers, including expert-novice comparisons. As Cross [20] points out in his summary, this work has revealed some interesting insights such as differences in strategies (structured organization of ideas and systematic expansion in experts versus exhaustive search in novices, or preliminary evaluation before detailed design in experts versus trial-and-error in novices) and that experts often conjecture solutions as a means to understand problems. However, this research is limited in a couple of ways. First, experts often exhibit behaviors that contradict what is often considered good design practice, for example not generating
multiple alternative ideas for problems that arise and a tendency to fixate on a given idea and persist with it even in the face of significant difficulties and shortcomings. Yet, we usually do not know whether experts could have done better had they approached the problem differently. That is to say, the design approaches acquired through experience may not be optimal. Also, as Cross points out, this research has led to little insight on how to help novices become expert designers.

In addition to prescriptive models and a growing body of descriptive research, quantitative techniques have been proposed to model and analyze task sequence and iterative sub-cycles within complex design projects. Smith and Morrow [24] identify a broad spectrum of modeling approaches including Design Structure Matrix [25,26], Signal Flow Graphs [27,28], queuing models, Markov chains, and scheduling models. These modeling approaches must assume a problem structure, and so are potentially useful for product development contexts where similar development projects occur repeatedly. They are less useful for our context of student design projects where each project is unique. Smith and Morrow [24] do not identify any statistical models based on empirical data, and in fact note that the biggest criticism of modeling efforts to date is their lack of applicability and usefulness.

Our intention, then, was to devise a study that would explicitly relate process to outcome and empirically validate several features of design processes that have been identified as important factors or are otherwise advocated in the literature. To do this, we create a computer model based on data from actual student design projects. We then perform a design of experiments using the computer model, the results of which provide
explicit links (include direction of effect and relative magnitudes) between process
variables and design outcomes.

This study makes three important contributions. First, we study actual design
processes of students engaged in mechanical engineering capstone projects. While we
found numerous studies using students as subjects, we found little design research
involving capstone projects particularly, despite their importance for accreditation in the
U.S. Second, we provide statistical validation for several factors associated with “good”
design processes consistent with Blessing, et al.’s [1] admonition for greater rigor in the
field of design research. The modeling and analysis results also suggest areas for further
investigation, including modifications to the conventional prescriptive models that may
help inexperienced (a.k.a., student) designers achieve more desirable end results. Third,
we describe a novel approach to studying design processes using computer design of
experiments that simulates real world design processes. Such an approach allows the
researcher to leverage powerful statistical techniques to identify patterns and relationships
among variables in the data where sample size is a limiting factor. The next section
describes our overall research design, including process data collection, quality
measurement, and analysis methods.

2 RESEARCH APPROACH

This study focuses on capstone mechanical engineering design projects completed
between Spring 2001 and Fall 2002 semesters at Montana State University (Bozeman,
Montana, USA). ME 404, Mechanical Engineering Design II, was a four-credit-hour one-
semester course at the time, taken in the senior year (it has since been converted to a two-
semester course). The course is preceded by ME 403, Mechanical Engineering Design I,
where students learn a design methodology and put it to use in a significant design project. In addition, students enrolled in ME 404 will have completed the junior year curriculum, and most will have participated in at least one summer internship. Thus our study group might be considered “semi-expert” designers [14]. Course enrollment reflects the student body in the major, with approximately 10% female representation and 5% international representation.

For the capstone project, students are divided into teams of two to four with a faculty member as advisor. The projects are industry sponsored so each team must interact with a client/sponsor to define their needs, devise a solution to meet those needs, and deliver a product (set of engineering drawings and specifications, written report, oral report, and in many cases a hardware prototype) by semester’s end. Each project is unique.

2.1 Data Collection for Process Variables

Researchers have used a number of techniques to collect data on design processes, including interviews [29,30], retrospective and depositional methods [31], protocol analysis [32,33] and process observation [34]. However, for this study, in order to study design processes *in-situ*, spread over 15 weeks (one semester), without a specified location for the work, while still capturing details of design work when and as they occur, each of these techniques seemed problematic. Thus, design journals seemed a logical choice.

**Design Journals.** Ball, et al. [14] used diaries as their primary data collection mechanism; however, unlike their highly structured approach of asking students to answer a specific set of questions weekly, we chose a less structured approach to avoid bias in the data collected and allow students to record what information they thought was important for their project. Students were required to keep individual design journals to document
their work over the semester as a part of their work package for the semester project [35]. Journals constituted 15% of the final course grade. At project completion, journals were collected, and a subset of the projects was coded for analysis.

Using student journals to collect data overcomes some of the drawbacks of other research methods. Compared to interviews, retrospective, and depositional methods, the data is collected in real-time, but unlike observational approaches, our method does not require specially trained professionals and avoids the possibility of artificially altering student behavior by having an observer present. Like protocol analysis, the data can be readily quantified using a suitable coding scheme, but it requires little researcher intervention during data collection and data is collected from processes in situ rather than in a laboratory setting. It is also more feasible to collect a relatively large sample size compared to videotaping or other approaches because the quantity of data captured, while still large, is more manageable.

As with any data collection method, disadvantages to the technique exist. Journals may offer an incomplete record of the design process. Where designers either keep imperfect records or are unaware of important information, the journals may fail to capture critical details regarding the development of the design project. Students may “backfill,” that is, record events in retrospect rather than as they occur, which can lead to omissions of key information. Training and feedback during the project can help students improve their record keeping skills and discipline, but ultimately an accurate journal record depends on the designer’s commitment to keeping a good journal.

We took a number of measures to help ensure the validity of the data collected and subsequently used in modeling and analysis. First, we provided in-class training during the
first week of classes on keeping a design journal, along with multiple reasons why this is a useful practice that they could and probably should carry forward into their professional careers, in an attempt to motivate thorough and consistent journaling. They were told to record “everything” concerning their projects, with specific suggestions on how to document certain common activities (e.g., meetings, brainstorming sessions, internet searches). However, we did not specify a specific format or content, save for the requirement to record dates, times, and attendees, leaving it up to the student to decide what to document.

Next we collected the student journals every 2-3 weeks over the course of the project, and evaluated them using a rubric designed to assess the thoroughness of the design journals (students received copies of the rubric on the first day of class). They were collected at 5:00 p.m. one day, and returned by 8:00 or 10:00 a.m. the following day so as to minimize the amount time students went without their journals. Students were required to put time stamps on journal entries, which enabled us to assess thoroughness more easily—obviously the expectations for journal content for a 15 minute meeting versus a 2 hour meeting are quite different. Students received specific feedback on how to improve their journal records via sticky notes in the journal (so the feedback could be made specific to an entry, but also could be removed), and they received a numerical score. Rubric scores tended to increase over the first two checks, then level-off at a moderately-high level until the end, where a slight dip would occur at the end-of-semester crunch. The ME faculty consensus among the more senior members was that they had never seen this high level of quality journals.
Furthermore, we observed weekly advisor meetings to gain a real-time snap shot into the project each week. From the meetings it was clear whether the students were using the journals, and whether the journal record represented actual project activity. Logistically it was not possible to observe every team in our sample, but we did observe about half of them and some others not in the sample. The journal records were remarkably consistent with the observations of the weekly advisor meetings.

In the end, despite our best efforts, not all journals were useable for this study due to journal quality. Thus, we selectively sampled the projects, choosing from among those projects with usable journal records only and screening those journals with significant backfill or gaps in documentation.

Finally to improve the representativeness of the data, we aggregate the data from the 3-4 members of each design team to the project level. Data from the separate records within each project corroborate one another in such meaningful and complex ways that collusion among the team members seemed a remote possibility (although we could identify a few instances where that apparently happened). A typical project has 300-600 pages of documentation depending on the size of the team, the number hours dedicated to the project, and their prolificacy. Aggregating this amount of data makes us reasonably confident the data is fairly representative of the actual processes used.

**Coding.** At project completion, journals were collected and coded according to the scheme in Table 1, with times assigned according to the start and end times recorded. Each design related activity received two codes, one for design activity and one for design level.
The design activity code delineates four categories along the lines of several factors considered important for design success. As described in the previous section, a number of researchers observe that problem scoping, problem setting, information gathering and related activities seem important. The problem definition (PD) code identifies when students gather or synthesize information to better understand a problem or design idea through activities such as: stating a problem, identifying deliverables, and researching existing technologies. A number of authors claim that generating multiple alternatives is another key factor, yet other studies indicate such a strategy is not pervasive in actual design processes. Thus, an idea generation (IG) code was defined to capture activities where teams explore qualitatively different approaches to recognized problems, such as brainstorming activities, listing of alternatives, and recording “breakthrough” ideas. Engineering curricula tend to emphasize engineering analysis, and several studies point toward solution analysis as key success factors. Thus an engineering analysis (EA) code is defined as formal and informal evaluation of existing design/idea(s), e.g., mathematical modeling and decision matrices. Finally, iteration and sequential solutions search strategies are mentioned in both the descriptive and prescriptive literatures. The design refinement (DR) code captures this, and includes activities related to modifying or adding detail to existing designs or ideas, deciding parameter values, drawing completed sketches of a design, and creating engineering drawings using computer-aided design (CAD) software.

The design level code was assigned one of three levels. Concept design (C) addresses a problem or sub-problem with preliminary ideas, strategies, and/or approaches. Common concept design activities are identifying customer needs, establishing the design specifications, and generating and selecting concepts. System level design (S) defines the
needed subsystems, their configuration and their interfaces (also referred to as product architecture design). Detail design activities (D) focus on quantifying specific features required to realize a particular concept, for example defining part geometry, choosing materials, or assigning tolerances. We also considered work done at the component level (e.g., fasteners, batteries, valves) detail level design. These levels are quite consistent with the prescriptive literature where three phases of design are often defined (e.g., concept, system-level, and detail design in Ulrich and Eppinger [36]; concept, preliminary, and detail design in Dym and Little [7]; or concept, embodiment, and detail design in Pahl and Beitz [9]). Ullman, Dietterich, and Stauffer also use a similar construct in distinguishing levels of design tasks in their task-episode accumulation model [37].

The dual code system (design level/design activity) allows us to distinguish between, for example, concept level design refinement and detail level idea generation, or system level engineering analysis and detail level problem definition. The coding scheme also designates codes for activities associated with project management (PM), report writing (RW) and presentation preparation (PP) so that every entry could be assigned a code. However, this study focuses only on the design activities described in the previous two paragraphs.

To give an idea of the kind of data captured in the student journals, a sample journal entry appears in Figure 1. The project was to design an automated chamomile flower harvesting device as a retrofit attachment to the client’s tractor-combine. In this entry, the student documents reading a hydraulics manual, apparently to learn how hydraulics might be used as the power source for their device (the tractor-combine has a hydraulic motor output for powering external attachments). He then notes where hydraulic
power might be useful in their evolving design concept. Thus, these first lines are all coded “C/PD” (concept level problem definition) as the activity appears to be gathering information about how hydraulic motors work and can be used followed by defining where/how for this application (problem definition), all at fairly preliminary (i.e., conceptual) level. Starting with the light bulb icon, it seems the student transitions to thinking about how the various components of the idea would work together, eventually drawing a sketch to visualize the system (system level). And since this is the first occurrence of this configuration in any of the project’s journals (i.e., a new idea), the sketch and preceding note receives an “S/IG” (system level idea generation) code. The next-to-last notation (sentence starting, “Need to…”) seems to be defining a problem to be addressed with this configuration, and so received an “S/PD” (system level problem definition) code. The student notes that this activity occurs over a 1.5-hour span, so each code is allocated a portion of the 1.5 hours according to the space apportioned in the journal entry (i.e., 0.50 hours of C/PD, 0.75 hours of S/IG, and 0.25 hours of S/PD).

The process of journal coding proceeded in two stages. First, research assistants familiarized themselves with the projects by reading the final written reports, then coded entries and captured times by walking through team members’ journals in lock step, considering all the members’ entries for a given day before assigning codes and times, then moving to the next day. Simple rules were devised for allocating time, and for resolving discrepancies among the different journal accounts. The principal investigator then reviewed the coding as a crosscheck on accuracy and consistency. Disagreements were resolved through discussion and the process continued until mutual agreement was
reached. The time data on the various process variables was then entered into a database, and aggregated for the project by combining individual journal data.

The sample size for this study is 14 design projects (47 100-page design journals). The 12 design variables (four activity categories across three abstraction levels) serve as the independent variables in the model constructed, with the cumulative times for each project, expressed as a proportion of total design hours, used as the variable values (see Sobek [38] for more journal coding details).

2.2 Design Quality Measurement

To measure the “goodness” of the students’ end products, we developed two outcomes measures, a client satisfaction score and a design quality index. Consequently, two separate instruments, the Client Satisfaction Questionnaire and the Design Quality Rubric (DQR), were developed, validated and deployed for measuring the client satisfaction and design quality quantitatively. In this paper, we report the methods and results associated with the DQR.

To develop the DQR, we first obtained evaluation schemes from mechanical engineering capstone course instructors at 30 top ranking mechanical engineering programs in the US. We also collected evaluation schemes from several design contests [39,40,41,42]. From the evaluation schema, we extracted 23 metrics common to the evaluation schema used to rate a design project. A number of the 23 metrics were similar enough (e.g., creativity, innovation, and originality) to aggregate into six broad categories: requirements, feasibility, creativity, simplicity, aesthetics and professionalism. Since aesthetics is not a requirement in many of our projects, and professionalism deals more with report/presentation quality attributes than design quality directly, we further reduced
the above to five metrics, replacing the last two with an “overall impression” question to capture the reviewer’s overall assessment, which could include professionalism and aesthetics if appropriate. The final metrics and their definitions are presented in Table 2. A seven-point scale was used for each question/metric and three anchors provided (1-poor, 4-acceptable, 7-outstanding). A brief rationale was requested from each evaluator on each response for the purpose of inter-reviewer comparisons to evaluate consistency among the evaluators.

Four engineering professionals were hired to evaluate the design projects. Three were licensed professional engineers, each with over 10 years of experience in design and manufacturing. The fourth had five years of experience and was not yet professionally licensed. These evaluators were asked to evaluate the project outcomes as presented in the students’ final reports. Specific instructions were provided to assess the design projects on the final products only (not on the process) as if they were evaluating actual industry designs, taking the project time and budgetary constraints into account. Each evaluator was assigned reports in such a way that each report was evaluated at least twice to provide redundancy in the measurement. All four evaluators looked at two of the reports in order to determine inter-evaluator consistency. The quality index for each project was calculated by averaging the scores of the individual metrics, then averaging across evaluators. Thus, the quality score is on a scale of 1-7. More detail on the development of the DQR can be obtained from Jain [43] or Sobek and Jain [44].

2.3 Data Analysis Approach

Design of experiments (DOE) is a powerful and efficient strategy employing statistical principles commonly used for experimental optimization and discovery of
important factors influencing product and process quality. However, the small sample size and high dimensionality of the data in this study pose significant challenges in using this technique. To address this concern, we first we constructed a mathematical model from the design journal data to simulate actual student design processes. Then, using this model as a base, we constructed a second model based on classical DOE strategy and principles. The second model is considered a metamodel; that is, a higher level model formed from an analytical, neural network, or computer model [45]. The metamodel enables us to study the process under desired conditions and gain insight into cause and effect relationships within the system without running an actual experiment. Since the basic approach uses a virtual experiment from which to draw conclusions, some researchers have termed it a virtual design of experiment (or VDOE) [46].

We modeled the journal data using a principal component neural network, a special class of neural networks designed for data with high dimensionality [47,48]. This hybrid architecture reduces the dimensionality of the data to help compensate for the small sample size, and allows output prediction in terms of the original variables. This base model was constructed using the percent of design time spent at each design level/activity combination per project (from Table 1) as input variables, and the design quality index (from the DQR) as the target variable. A subset of the sample was used to train the networks and the remaining sample was used to cross-validate. Several different network architectures were constructed and trained using Neurosolutions software [49]. The best network was chosen using the mean squared errors (MSE) of the training and cross validation sets as the judging criteria.
To determine the relationships among the design process variables and the outcome measure, we analyzed a $2^{12-4}$ fractional factorial with design quality as the response variable. The data for each run in the design grid for the VDOE metamodel was obtained from the artificial neural network model [50,51,52]. Due to the deterministic nature of the neural network base model, classical notions of experimental unit, blocking, replication and randomization are irrelevant in the VDOE.

The final factorial was a resolution V design with 299 runs. Data transformation, model fitting, analysis of variance (ANOVA), model reduction and model adequacy checking were all performed in Design Expert software [53] to obtain the response curves for various factors and factor interactions. Response was predicted under various process settings within the range of the data utilized to construct the model (see Jain [43] for more detailed information on neural network / VDOE methodology).

3 RESULTS

Table 3 reports the means and standard deviations of the process and outcome data used in the modeling. The values for the process variables are expressed as percent of total design time. For example, teams in the sample spent an average of 13.14% of their total design effort (as reported in the journals) on concept level problem definition activity. The number of hours varies considerably across the design process variables with the most time spent on design refinement work. A correlation analysis of the 12 variables found only 2 pairs of variables out of a possible 72 to be significantly correlated at a 1% significance level.

Table 4 presents the architecture summary of the neural network model constructed. The principal components network reduced the original 12 variables to six
independent components explaining 99% of the variation in the data. The best performing network (based on the judging criteria) contains a single hidden layer with 2 hidden neurons. From the learning results, it was observed that the established network architecture had a good “memory” and the trained matrices of weights and bias reflected the hidden functional relationship well. Thus the model can serve as a reasonable surrogate to reality. Finally, because the testing and validation MSE were small, the models developed can be considered reliable for the prediction of the response scores under any combination of the process parameters as long as they are within the ranges investigated.

Next, Table 5 presents the analysis of variance (ANOVA) results for the experimental design results. The insignificant factors are not included at a significance level of $p \leq 0.05$. The large values of the F-ratios and small p-values suggest that the terms significantly affect the response. Interactions between the individual variables follow the same trend as their primary effect, except for C/EA and D/EA which are insignificant as primary effects but are significant in interactions.

To estimate the relative importance of the significant factors in the VDOE model, the slopes of each variable versus the response variables were taken from the response plots of the ANOVA, then divided by the absolute value of the smallest magnitude slope among significant variables (i.e., D/DR). Table 6 reports these figures, and indicates that the strongest positive effect is that of system level engineering analysis (its effect is approximately 100 times stronger than that of detail level design refinement). Concept and system level problem definition, and system level idea generation activities also have significantly positive effects. On the negative side, conceptual level design refinement and
The data used in modeling is normalized by total design time. This proportionality implies time spent doing X is time spent not doing Y, which clearly affects the interpretation of the results. Four trends can be identified from the results reported in Table 6:

1. Proportionately more time spent on system level design has the biggest positive association with design quality, particularly problem definition, idea generation, and engineering analysis activities.

2. Problem definition (PD) at the higher abstraction levels appears positively related to design quality.

3. Idea generation (IG) at concept level is negatively associated with design quality, but idea generation at the system level is positively associated.

4. Proportionately greater time spent on design refinement (DR) activities or at the detail level is negatively associated with design quality or is insignificant.

We develop these themes in more depth in the following subsections, discussing them in light of the extant literature presented in Section 1, and then finish the discussion with study limitations. In particular, we indicate where our results provide validation for previous findings, and where they point towards further research.
4.1 System Design Work

The most striking trend in Table 6 is the highly positive effects of system-level design work, particularly problem definition, idea generation, and engineering analysis activities. Given that the student teams in our sample spent only about 9% of their design time on average in system-level work (see Table 3), this result suggests that system-level design is a high-leverage activity. What we call system-level design is similar to Ulrich and Eppinger’s system-level design [36] as well as what Ullman calls configuration design [11], Dym and Little call preliminary design [7], and Pahl and Beitz call layout design [9].

Conceptual ideas are difficult to evaluate, particularly for inexperienced designers. But by fleshing out the system-level design—that is, understanding subsystems and how they interact, and exploring alternative configurations of components and subsystems—the design team can get a much better estimate of the performance of an idea without spending the many hours it takes to detail a design. Adjustments at the system level are fairly easy to make, while adjustments at the detailed level (e.g., in a set of detailed CAD drawings or prototype) are comparatively time consuming. So it seems that effort levied at system-level issues can prevent time-consuming adjustments later in the design process.

Perhaps just as importantly, the system architecture (or configuration or layout) might well be considered a design problem in its own right. Even after concept selection, there is usually considerable leeway in how that concept can be configured. Much may be gained by repeating the proper design steps on this important sub-problem, e.g., problem definition and information gathering, search for alternative solutions, and so forth. The tendency among students is to envision a concept in a particular configuration and rarely explore alternative configurations that could turn a ho-hum design concept into a clever,
innovative design. The data indicate that teams which bucked this tendency and did some (even if limited) system-level design achieved better design quality. Thus, lack of system-level design may be costly on two accounts: failing to explore all the opportunities afforded a given a concept, and lesser ability to distinguish less from more problematic concepts.

While a number of the design process models in the prescriptive literature skip over this intermediate level, a number of them emphasize some sort of intermediate stage or step between concept design and detailed design. Digging deeper, though, we find that comparatively little guidance and few tools can be found to help with this intermediate stage of design work.

The descriptive research seems to have not yet addressed system level design work much. We found only a couple of studies. Ahmed, Wallace and Blessing [54], in their study of differences in design patterns between novice and more experienced designers, found that the novice design pattern was to generate ideas, implement them, and then evaluate. Whereas, experienced designers tend to add a fourth step, “preliminary evaluation”, between generate ideas and implementation. Our results extend this finding by providing some validation that the “added step” improves design quality, even among inexperienced designers.

Newstetter and McCracken [55] found that student designers tend to jump from conceptual to detail-level work, skipping intermediate-level work. Ignoring this step leads to a higher probability that the design will have to be revised, thereby leading to a trial and error pattern. Our results provide quantitative evidence that skipping intermediate-level work leads to lower quality designs.
These results suggest that, given the apparent high impact on design quality and that few tools are currently available to aid designers in system-level work, more research is warranted in how to transition from conceptual design thinking to the detailed design phases. Additional work to better understand system-level design appears to hold significant potential for increasing the productivity of design engineers.

4.2 Problem Definition and Scoping

The analysis results indicate that proportionately greater time spent on problem definition (PD) activity at the higher abstraction levels is positively associated with design quality. Many of the activities coded “PD” are information gathering, while others are sense-making activities on the collected information. Conceptual level problem definition includes activities like an internet or library search on existing design solutions, interacting with the client to clarify the problem space, researching basic design mechanisms or analysis methods, and examining existing designs of others. Similarly, system level problem definition (as seen in design journals) includes activities like exploring requirements for the various subsystems, identifying the constraints on interfacing mechanisms, and understanding the final assembly sequence for the design. Interestingly, the effect of similar activities at the detailed design level is negative.

These results confirm the findings of Atman, et al.’s [17] comparison of university freshmen and senior design processes and Adams, et al.’s [18] summary of several empirical studies on student design processes. They found that problem scoping cycles and problem setting activities have a positive association among first-year and senior engineering students designing in a laboratory setting. Our results extend these findings to actual student capstone design projects, but only for higher abstraction levels. They also
provide additional, though weaker, confirmation of studies that found goal analysis as important to design success [15,19].

A number of studies have found that it is possible to spend too much time in problem definition, information gathering, and related activities [17,19,20], which has a negative effect. It is possible that the negative association between PD activities at the detailed level reflects excessive time spent in these kinds of activities. But more investigation is needed.

4.3 Idea Generation

Another intriguing result of our analysis concerns the effect of idea generation (IG) in the sample. Even though it is a generally accepted precept in the prescriptive literature that good designs result from processes that consider multiple alternative solutions, and that some descriptive research supports this claim, a number of studies have found that multiple competing alternatives is a practice rarely found in actual design processes. Our results seem to reflect these mixed findings, but with a twist. Proportionately greater time spent on idea generation (IG) at the concept design level is strongly negatively related to design quality, while idea generation at the system design level shows a strongly positive association with design quality. Idea generation at the detail design level is insignificant. One might expect idea generation at all levels to track together, but this was not the case in this study.

These results might be explained in part by another study. Newstetter and McCracken [55] found that engineering students perceive design as generating lots of ideas without much consideration to the merit of the ideas, a misconception the authors term “ideation without substance.” Within our data set, we found that teams often include
alternatives in the final report that were not given serious consideration, or spend time brainstorming alternative design ideas simply because their faculty advisor requires it. The alternative ideas are often seen as throw-away ideas; that is, not considered serious contenders but rather created to satisfy their advisor or to show that they considered their options. This shows up in our analysis as a negative effect of conceptual idea generation on design quality. However, idea generation at the system level seems to reflect a greater seriousness about the alternatives they generate, and so associates positively with design quality.

This finding has important implications for design educators and for newly hired engineers. While it may seem contrary to popular views of design, we hypothesize that students are better off refraining from spending excessive time on idea generation/brainstorming activities. Rather, in light of the effect of problem definition discussed above, inexperienced designers may achieve better outcomes by spending more time researching existing technology and design solutions, and let ideation flow from there. Testing this assertion is the subject of future work.

4.4 Detail Level Work and Design Refinement

As the last trend we will discuss, Table 6 shows that detail level work is negatively associated with design quality or is insignificant across activity types. Similarly, the effect of design refinement (DR) activity is negative on two design levels and insignificant on the third. Design refinement activities are those that modify existing ideas and design solutions and/or that add the finishing details on designs (e.g., specifying tolerances or fasteners). Most CAD work, prototyping work, and design changes based on test or analysis results are considered DR. Student teams spend an average of 70% of their total
design time on detail design level work, and 40% of total design time on design refinement activities. The VDOE results imply that design teams achieve higher quality scores with greater emphasis placed on concept level problem definition and system level design activities, so that detailed level and design refinement work require less effort.

This implication is consistent with authors on design who agree that the early stages in the design process are the most important. One of the possible reasons for this could be that those design teams which skimp on the conceptual and system level design must compensate for it with additional detail level and/or design refinement work, and the trade is not one-for-one. This further suggests that there are diminishing returns associated with the different levels of design abstraction. The incremental benefit of effort spent at higher levels of abstraction is comparatively greater than the incremental benefit of detailed design work.

Interestingly, these results do not confirm a finding of some other studies that associate solution analysis with successful design outcomes\textsuperscript{15,19}. Only engineering analysis at the system-level appears positively associated with design quality in this study; analysis work at conceptual and detailed levels has no association. However, that detailed engineering analysis is insignificant does not necessarily contradict the findings of these other studies. It simply means these activities did not distinguish the projects within our sample. All design teams devoted significant effort to detailed engineering analysis, so we have no data from teams that did very little solution analysis.

4.5 Limitations & Future Continuations

Like most, this study has its limitations. The sample size used to draw the conclusions may be an issue as small sample sizes can produce inaccurate or misleading
results. However, the data used are aggregate measures and potentially possess strong explanatory power. For example, the 14 projects in the sample represent 47 student journals, over 4,000 pages of documentation, and some 8,600 person-hours of effort. Since journal data were aggregated to the project level, each data point represents dozens if not hundreds of person-hours. So each data point is fairly robust, diminishing the drawback of a small (in statistical terms) sample size.

Second, the results apply only to the range of the data in the sample. Thus, for example, that detail level design refinement is negatively associated with design quality does not mean that student design teams should simply stop doing these activities. Rather, because zero is not within the range of the data in the sample, this result simply means that student design teams should strive to structure their design projects in such a way as to reduce the amount of time and effort required in these activities. The conclusions cannot be extrapolated beyond the range of the sample data.

Next, the rubric used to measure design quality may include bias despite careful measures to avoid it. Furthermore, the data collected in this study (both process and outcome) is to some extent subjective. It can also be argued that the data collected from design journals can be inaccurate, incomplete or biased. We addressed these limitations through a rigorous cross-checking procedure of the journal coding, redundant evaluations, and using neural networks which are designed specially for noisy data. Still, more studies should be conducted to substantiate these findings.

Another limitation of this study are the variables not considered, such as effort in “non-design” activities, personality types, team diversity, advisor effects, team experience, and project-related characteristics (e.g., whether a prototype was required, whether it was a
“clean sheet” project or not). Some may see this as a limitation as these could have provided more insight into the results. But in some ways it actually strengthens the study: we get significant results without accounting for all of these other sources of variability! The effects of process, therefore, must be fairly strong.

Lastly, the sequence and timing of the occurrences of the various process variables was not considered. It is possible that the timing of the various activities is just as important as whether they occur or not and in what amounts. Thus future work will seek to identify the significance of the sequence of the various design process variables.

5 CONCLUSION

This study attempted to gain insight into design processes by studying how process variables affect outcomes in student engineering projects. We collected data from 14 projects (representing 47 students in total) and modeled the data using an artificial neural network. Then by performing a virtual design of experiments, using the artificial neural network model to predict the magnitude of the response variable, we were able to obtain estimates of the relative impacts of the 12 design process variables used. In other words, we could answer which process variables are positively or negatively associated with project outcomes, and the relative magnitude of those effects. The study fits in Blessing, et al.’s general design research methodology by trying to link the descriptive and prescriptive realms, and providing some level of validation for factors believed to influence the success of design projects.

Specifically, the results support the propositions that problem definition is important to design quality, that earlier design phases have comparatively greater impact, and that intermediate design levels falling between concept and detailed design are
important. Results regarding idea generation reflect the mixed results in the literature, so more investigation is required.

In addition, our results suggest modifications to commonly accepted process models to make them more applicable to inexperienced designers. Numerous studies have found significant differences between novice and expert designers across varied fields of study\textsuperscript{20,54,56,57}. Such research begs an important question: how can one design process model be well-suited for both novice and expert designers?

Our study suggests that engineering design process models can be modified in at least two respects to produce better design outcomes for engineering students and other inexperienced designers. First, problem definition and information gathering activities should receive greater prominence than an obligatory mention and exhortation that “it’s important” for good design. Novice engineers (e.g., students) need greater guidance on how to identify and frame problems, what kinds of information to gather, and how to organize and use it. Further, our results suggest that novice engineers should not be encouraged to “try to come up with some ideas,” advice commonly heard from advisors. Rather, they should be encouraged to research existing solutions to similar or analogous problems. In doing so, and in trying to improve them, the novice engineer begins to build that experience base that will enable him/her to become an expert designer.

Secondly, our results strongly suggest that students should be encouraged to delay jumping to detailed design until sufficient system-level work has been done. This could be another way to avoid ideation without substance—require students to flesh out and evaluate any idea at the system level before considering it a bona-fide alternative, which further implies doing system level design work before concept selection. Plus, an ability to
define a given solution at the system level may be a reliable rule-of-thumb test of whether sufficient problem definition work has been done. The challenge here is that system-level design tools are still rare. Additional work is currently underway to flesh out these two modifications and test them.

We concur with Blessing, et al.\textsuperscript{1} that more work should be done to substantiate and validate design models and tools. This study is one step in that direction. Of course, these results are not conclusive (sample consists of 14 projects from one discipline at one institution), but they begin to lay the groundwork for additional studies to substantiate advice and findings from the design literature, something that, to our knowledge, has not be done using statistical analysis. In addition, research into how design/engineering expertise is acquired would be highly beneficial. As previously mentioned, the timing and sequence of design process steps is another possible avenue of investigation. New representations and tools for systems-level design and analysis are needed. In short, a good deal of work is still lies ahead before we fully understand how to help our students become the best design engineers they can be.

\textbf{Acknowledgements}

Funding for this work was provided by the National Science Foundation (Grant No. REC – 9984484). Special thanks to the ME404 students and instructors for their cooperation and interest in this project, and to Steve Angell, Brian Berg, Grey Lerner, Robert Lowis, Clark Mozer, Thaddeus Orosz, Seth Partain, and Samuel Wilkening for their assistance.

2 Accreditation Board of Engineering and Technology


6 Cross, N Engineering Design Methods, John Wiley and Sons, Chichester, 1989.


30 Brockman J Evaluation of student design processes, Frontiers in Education
   Conference, Salt Lake City, 1996.


41 ASME Student Manufacturing Design Competition Rules

42 MHEFI—2003 Material Handling Student Design Competition Rules

43 Jain, V K Relating Design Process To Design Outcome In Engineering Capstone

44 Sobek, D K, and Jain, V K Two instruments for assessing design outcomes of
capstone projects, to appear in the Proceedings of the American Society of

45 Sacks, J, Welch, W, Mitchell, T, and Wynn, H Design and analysis of computer

46 Pyzdek, T Applying virtual DOE in the real world, Quality Digest (August 1999),

47 Diamantaras, K, and Kung, S Principal Component Neural Networks, John Wiley,
New York, 1996.

48 Tan, S, and Mavarovouniotis, M Reducing data dimensionality through optimizing

49 Neural Networks-Neurosolutions http://www.neurosolutions.com/, accessed June
2004.


List of Figures

Figure 1: Sample Journal Entry
List of Tables

Table 1: Coding Matrix
Table 2: Design Quality Rubric
Table 3: Summary Statistics
Table 4: Network Architecture
Table 5: ANOVA Results of the Design Quality Experimental Design Results
Table 6: Relative Factor Slope Scaling
Table 1: Coding Matrix

<table>
<thead>
<tr>
<th>Design Activities</th>
<th>Concept (C)</th>
<th>System (S)</th>
<th>Detail (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Definition (PD)</td>
<td>C/PD</td>
<td>S/PD</td>
<td>D/PD</td>
</tr>
<tr>
<td>Idea Generation (IG)</td>
<td>C/IG</td>
<td>S/IG</td>
<td>D/IG</td>
</tr>
<tr>
<td>Engineering Analysis (EA)</td>
<td>C/EA</td>
<td>S/EA</td>
<td>D/EA</td>
</tr>
<tr>
<td>Design Refinement (DR)</td>
<td>C/DR</td>
<td>S/DR</td>
<td>D/DR</td>
</tr>
</tbody>
</table>

<p>| Non-Design Activities                  |             |            |            |
| Project Management                     | PM          |            |            |
| Report Writing                         | RW          |            |            |
| Presentation Preparation               | PP          |            |            |</p>
<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic</strong></td>
<td></td>
</tr>
<tr>
<td>Requirements</td>
<td>The design meets the technical criteria and the customer requirements</td>
</tr>
<tr>
<td>Feasibility</td>
<td>The design is feasible in its application and fabrication / assembly</td>
</tr>
<tr>
<td><strong>Advanced</strong></td>
<td></td>
</tr>
<tr>
<td>Creativity</td>
<td>The design incorporates original and novel ideas, non-intuitive approaches or innovative solutions</td>
</tr>
<tr>
<td>Simplicity</td>
<td>The design is simple, avoiding any unnecessary sophistication and complexity, and hence is: Practical Usable, Reliable Ergonomic, Serviceable Safe</td>
</tr>
<tr>
<td>Overall</td>
<td>Overall impression of the design solution</td>
</tr>
</tbody>
</table>
Table 3: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean (%)</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Process Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/PD</td>
<td>13.14</td>
<td>9.28</td>
</tr>
<tr>
<td>S/PD</td>
<td>2.16</td>
<td>3.27</td>
</tr>
<tr>
<td>D/PD</td>
<td>8.68</td>
<td>6.10</td>
</tr>
<tr>
<td>C/IG</td>
<td>4.41</td>
<td>2.45</td>
</tr>
<tr>
<td>S/IG</td>
<td>2.83</td>
<td>1.90</td>
</tr>
<tr>
<td>D/IG</td>
<td>2.78</td>
<td>2.87</td>
</tr>
<tr>
<td>C/EA</td>
<td>2.94</td>
<td>3.82</td>
</tr>
<tr>
<td>S/EA</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td>D/EA</td>
<td>24.44</td>
<td>16.72</td>
</tr>
<tr>
<td>C/DR</td>
<td>1.39</td>
<td>2.55</td>
</tr>
<tr>
<td>S/DR</td>
<td>3.54</td>
<td>3.48</td>
</tr>
<tr>
<td>D/DR</td>
<td>32.93</td>
<td>16.90</td>
</tr>
<tr>
<td><strong>Outcome Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DQR</td>
<td>4.42</td>
<td>1.06</td>
</tr>
</tbody>
</table>
### Table 4: Network Architecture

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Quality Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of input variables</td>
<td>12</td>
</tr>
<tr>
<td>Number of principal components</td>
<td>6</td>
</tr>
<tr>
<td>Number of hidden layers</td>
<td>1</td>
</tr>
<tr>
<td>Number of hidden neurons</td>
<td>2</td>
</tr>
<tr>
<td>Training set</td>
<td>11</td>
</tr>
<tr>
<td>Cross-validation set</td>
<td>3</td>
</tr>
<tr>
<td>MSE (training set)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>MSE (cross-validation set)</td>
<td>&lt; 0.21</td>
</tr>
</tbody>
</table>
Table 5: ANOVA Results of the Design Quality Experimental Design Results

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>209.95</td>
<td>22</td>
<td>9.54</td>
<td>24.06</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>C/PD</td>
<td>3.11</td>
<td>1</td>
<td>3.11</td>
<td>7.84</td>
<td>0.0055</td>
</tr>
<tr>
<td>S/PD</td>
<td>40.97</td>
<td>1</td>
<td>40.97</td>
<td>103.32</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>D/PD</td>
<td>20.52</td>
<td>1</td>
<td>20.52</td>
<td>51.74</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>C/IG</td>
<td>22.86</td>
<td>1</td>
<td>22.86</td>
<td>57.63</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>S/IG</td>
<td>6.78</td>
<td>1</td>
<td>6.78</td>
<td>17.11</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>S/EA</td>
<td>22.72</td>
<td>1</td>
<td>22.72</td>
<td>57.28</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>C/DR</td>
<td>43.47</td>
<td>1</td>
<td>43.47</td>
<td>109.61</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>D/DR</td>
<td>1.78</td>
<td>1</td>
<td>1.78</td>
<td>4.50</td>
<td>0.0348</td>
</tr>
</tbody>
</table>
Table 6: Relative Factor Slope Scaling

<table>
<thead>
<tr>
<th></th>
<th>PD</th>
<th>IG</th>
<th>EA</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>+ 5.0</td>
<td>- 36.5</td>
<td>*</td>
<td>- 49.0</td>
</tr>
<tr>
<td>S</td>
<td>+ 40.5</td>
<td>+ 31.6</td>
<td>+ 114.5</td>
<td>*</td>
</tr>
<tr>
<td>D</td>
<td>- 14.8</td>
<td>*</td>
<td>*</td>
<td>- 1.0</td>
</tr>
</tbody>
</table>

* Insignificant at p ≤ 0.05