Evaluating the Impact of Teaching Function in an Engineering Design Curriculum

Dr. Robert L. Nagel, James Madison University

Dr. Robert Nagel is an Assistant Professor in the Department of Engineering at James Madison University. Dr. Nagel joined the James Madison University after completing his Ph.D. in mechanical engineering at Oregon State University. He has a B.S. from Trine University and a M.S. from the Missouri University of Science and Technology, both in mechanical engineering. Since joining James Madison University, Nagel has helped to develop and teach the six course engineering design sequence which represents the spine of the curriculum for the Department of Engineering. The research and teaching interests of Dr. Nagel tend to revolve around engineering design and engineering design education, and in particular, the design conceptualization phase of the design process. He has performed research with the US Army Chemical Corps, General Motors Research and Development Center, and the US Air Force Academy, and he has received grants from the NSF, the EPA, and General Motors Corporation.

Prof. Matt Robert Bohm, University of Louisville
Dr. Julie S Linsey, Georgia Institute of Technology

Dr. Julie S. Linsey is an Associate Professor in the George W. Woodruff School of Mechanical Engineering at the Georgia Institute of Technology. Dr. Linsey received her Ph.D. in Mechanical Engineering at The University of Texas. Her research area is design cognition including systematic methods and tools for innovative design with a particular focus on concept generation and design-by-analogy. Her research seeks to understand designers’ cognitive processes with the goal of creating better tools and approaches to enhance engineering design. She has authored over 100 technical publications including over thirty journal papers, five book chapters, and she holds two patents.
Evaluating the Impact of Teaching Function in an Engineering Design Curriculum

Abstract
Functional modeling is often covered as a critical element of the engineering design process in engineering design texts, but little empirical data clearly demonstrates that functional modeling improves engineering designs or that teaching functional modeling makes students better designers. The overall objective of this project is to determine the impact of teaching function on engineering students’ design synthesis abilities. Two studies are being performed as a part of this project: (1) a longitudinal study following students through their sophomore, junior, and senior year following some being taught functional modeling, while others not, and (2) a yearly study looking at capstone project quality of students from cohorts either taught or not taught functional modeling. This paper focuses on preliminary data collected as a part of the longitudinal study using a functional modeling skills quiz to assess students’ ability to understand and represent a system. In particular, a functional modeling skill assessment quiz is being investigated for its ability to discern the extent of a student’s function knowledge. Two student groups are studied, one taught functional modeling along with function enumeration, and a second taught only function enumeration. The results provide promise that the skills quiz is working as desired; however, work is yet needed to develop an adequate scoring technique.

1 Introduction
Many engineering design texts discuss and prescribe functional recognition and some form of modeling as a step in the engineering design process. More empirical data is needed to show that teaching students functional modeling improves their design skills. This inclusion of functional modeling in many engineering design textbooks seems to demonstrate a recognized importance that engineering students should be taught to understand and abstract systems using function, yet there is only anecdotal evidence that suggests the students who are taught functional modeling during the process become better designers. The overall objective of the project is to determine the impact of teaching function on engineering students’ design synthesis abilities. We seek to answer the next logical question: Do students who are taught functional modeling become better designers?

More specifically, this research focuses on investigating the relationships between functional modeling skill and design outcomes by measuring the ability to (1) explore the solution space during ideation, (2) generate high quality designs, and (3) represent and understand engineered systems. Three hypotheses are being tested: (1) Functional modeling skills increase the quantity, variety, quality, and novelty of design alternatives during the concept generation phase of a design task; (2) Functional modeling skills increase the quality of a final design in a capstone course as judged by a group of faculty and industry experts; and (3) Functional modeling skills increase students’ ability to understand and represent a system. This paper focuses on one key assignment used in this broader study. Specifically, the goal of this study is to ensure that the assignment being used in this study provides an adequate measure of ability to model and understand function.

Function is broadly taught as a tool for moving from customer speak into engineering terms and allowing an engineer to formulate a complex problem into components that are more easily
solvable. Due to this identified key role in the design process, the results of this research have the potential to directly impact the two ABET student outcomes most directly related to engineering design: (C) an ability to design a system, component, or process to meet desired needs within realistic constraints, and (E) an ability to identify, formulate, and solve engineering problems. Consequently, the results from this project will provide both local and global impact within the field of engineering education. To ensure more broadly applicable and transferable results, the impact of learning function is being studied at two different types of public universities—each with its own approach to teaching functional modeling—one has functional modeling spread throughout the curriculum from sophomore through senior year and the other only teaches functional modeling during the senior capstone course. Studies will be cross-sectional at the senior level to allow the PI-team to understand how differences in prior educational experiences impact the value of learning function as well as longitudinal to allow the PIs to study the long-term impact of learning function. This paper focuses on preliminary data collected as a part of the longitudinal study using a functional modeling skills quiz, to assess students’ ability to understand and represent a system.

2 Background

Across numerous domains (e.g., systems engineering, control theory, computer engineering, and engineering design) function is considered as a technique for engineering abstraction allowing for complex systems to be modeled and simplified into a form more readily solvable. A review of literature of engineering design texts reveals a variety of techniques for teaching (and consequently, for representing) system functionality briefly reproduced below.  

- Glass Box Method: questions are asked about the design to move from a black box to a transparent box where additional functionality can be identified.
- Systematic Processes: a collection of methodological approaches that stem most directly from the work of Pahl and Beitz.
- Enumeration: meaning to “mention separately as if in counting; name one by one; specify, as in a list,” is the listing of each function-flow pair required by a system.
- Function (Means) Trees: fall under the category of hierarchical modeling approaches.
- Reverse Engineering: is performed while the modeler is in possession of the physical artifact allowing decomposition for understanding of the functionality.

Models tend to be based on “what a system should do” and are based on customer needs, design objectives, specifications, constraints, etc. Some models, such as function (means) trees may include solutions, while others such as flow-based functional models (such as the example provided as Figure 1) tend to be solution independent. In this research, we have chosen to focus on flow-based functional models stemming from the Pahl and Beitz methodology as they are perhaps one of the most common forms of functional models in engineering design. In addition, with the use of the Functional Basis, they have also been shown to be more repeatable than a prior flow-based approach.
Flow-based functional models tend to be hierarchical with two levels of abstraction—a black box model and a sub-functional model—linked by flows of materials, energies, and signals. Black box functional models are stand alone functional models abstracting a high-level transformation intended for the product to complete and are generated based on the system design requirements. The black box model describes the high-level transformation intended for systems, and the input and output flows identify all flows required for the operation of the product. A functional model decomposes the overall functional black box into specific flow transformations. Flow transformations define the operations required of the system such that the identified input flows do become the identified output flows through the operation of the system. Material flows are bold arrows; energy are thin arrows and signals are dashed arrows.

The authors have performed studies to assess the quality of student generated functional models using various methods of teaching and presenting functionality. Studies compared the effectiveness of three different techniques of presenting function: (Technique 1) lecture, in-class examples, and course text book, (Technique 2) a step-by-step example with black box model and functional model for a product in addition to everything in Technique 1, and (Technique 3) an algorithmic approach and grammar rules describing how function-flow pairs connect to form a functional model in addition to everything in Technique 2. Student generated functional models were evaluated using an 18 question scoring rubric. Each question was judged using a binary scale where the students were either correct (1), or incorrect (0). The scoring rubric focuses on assessing the mechanics of functional models but not the process of generating or the value derived from a functional abstraction. Prior to use of the scoring rubric, an inter-rater agreement (Pearson’s correlation) of 0.92 was obtained for the total score each student received. Findings of the studies show that thorough step-by-step examples led the students to generate better functional models, Technique 2. The additional grammar rules, Technique 3,
did not provide additional benefits. These studies informed the methodology applied to the study described herein.

3 Methodology

Preliminary data collected from a between-subject controlled experiment evaluates the students’ functional modeling skills and ability to represent a system functionally using the Functional Modeling Skill (FunSkill) test. The FunSkill test was developed by Linsey et al.\textsuperscript{29} as a tool to determine the impact of using high complexity design problems versus lower complexity design problems on functional modeling skills. It has been implemented on a range of studies with undergraduates having no functional modeling experience to graduate students having more extensive training in functional modeling. The skill test is a combination of open-ended questions with standardized correct answers and multiple-choice questions. It has face validity and produces adequate variability in senior undergraduates without showing ceiling or floor effects. For this study, questions focusing on function means trees models were removed from the original FunSkill test as students were not taught this form of modeling. The FunSkill test contained four questions: (1) identification of function statements from specifications, objectives, functions, and other items commonly confused with functions, (2) enumeration of potential function statements for a system, (3) enumeration of functions from design objectives, and (4) generation of a functional model. Questions were modified slightly for the students participating in this study such that specification, objective, and function wordings aligned with the course text (\textit{Engineering Design: A Project-based Introduction}). Questions are provided as Figures 2 through 5.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
\textbf{Is the following a function?} & \textbf{Yes} & \textbf{No} \\
\hline
weighs less than 200 lbs & & \\
store object & & \\
convert human energy to rotational energy & & \\
cut well & & \\
device will be easy to open / close & & \\
file finger nail & & \\
orient an object & & \\
convert force & & \\
easy to crank & & \\
cost less than $500 & & \\
\hline
\end{tabular}
\caption{Table 1. Identification of Function Statements from Specifications, Objectives, Functions, and Phrases often confused with Functions}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Figure 2. Question 1, Identification of Function Statements from Specifications, Objectives, Functions, and Phrases often confused with Functions}
\end{figure}
Questions 2, 3, and 4 required function enumeration and/or modeling at multiple levels of abstraction allowing for assessment of not only ability to identify, enumerate, and model function, but also to understand function as a hierarchical system representation.
Two conditions are compared: (1) an experimental group taught to enumerate function as well as model function using both black box and sub-functional modeling levels during their design course and (2) a control group taught only to enumerate function. The hypotheses are that:

1. The control and experimental group will equally be able to identify function statements,
2. The experimental group will perform better than the control group at enumerating sub-function-level function statements but no different at enumerating black box-level statements, and
3. The control group will be unable to generate black box or sub-functional functional models.

Based on the results of prior studies,\textsuperscript{25-28} students in the experimental group were provided with readings and taught during class using a combination of lecture and active learning examples, and were provided with detailed step-by-step examples that show how a black box model and a functional model are generated following the modeling steps taught during class. For the step-by-step example, students were provided the following four steps annotated with an example black box and/or functional model developed to a level appropriate for each step.

1. Generate a black box model for the product being designed considering the input flows, output flows, and the overall functionality of the product. Flows and function should be identified from the customer needs for the product.
2. Follow the Functional Basis \textsuperscript{30} or similar approach for the generation of the function-flow pairs. Generate function chains for flows identified at the black box level. Follow the material, energy, and signal flow convention from Step 1 when generating function chains. Add flows to represent the importation and exportation of materials, energies, and signals into the functional system.
3. Aggregate function chains. Add flows to represent the importation and exportation of materials, energies, and signals into the functional system.
4. Verify that all input functions identified in the black box model transition to all output flows in the black box model.

Students generated functional models first independently as a homework assignment (provided in Nagel et al.\textsuperscript{28}). Feedback was provided by the course instructor on student submissions. Students then worked as teams to generate a second functional model for the course project. The FunSkill quiz was given to both the control and the experimental group six weeks following the initial instruction of functional modeling in the experimental group. The study described herein took place in an undergraduate sophomore design class at a regional university with an undergraduate-only engineering program. Consent was provided by 79 sophomore engineering students. All remaining sophomore engineering students were of similar background with their only prior engineering course having been two introductory engineering courses completed during the prior academic year. Students could choose to take either a 9:30AM or 12:05PM section of design on Tuesday or a 9:30AM or 12:05PM section of design on Thursday. Students’ choice in Mathematics and Science (Physics or Chemistry) class is the primary impact on their design course section. For example, if students make the choice for lunch or afternoon design, they take morning science or math. Based on student background and course enrollment logistics, there is no reason to believe that one class is comprised of stronger students than another.
Tuesday sections (Sections A & B) were assigned to the experimental group while Thursday sections (Sections C & D) were assigned to the control group.

4 Scoring Process & Results

Scoring occurred in three phases and was completed by individuals not affiliated with the University where the study was run. Scorers were not aware of the Section treatments. All statistical analysis was completed with IBM SPSS Statistics v22. First, Questions 1, 2, and 3 were scored for general correctness. In other words, did the student recognize that the statement was or was not a function or was the student able to articulate a function statement (as a verb-noun pair) when prompted. For Question 1, the highest score a student could achieve was a 10, and for Question 2 and 3, the highest score a student could achieve was a 4. This scoring was completed by two expert modelers at two different institutions, and scores were averaged together for analysis. The correlations between the scores given by the expert modelers follow: Pearson’s Q1 = 0.98, Cohen’s Kappa Q2 = 0.53 (Moderate agreement), and Cohen’s Kappa Q3 = 0.61 (Substantial Agreement).31

Second, Questions 2 and 3 were scored by the same expert modelers to identify high level (i.e., black box functions) and low level (i.e., functional model functions). Scoring involved counting the number of high level functions articulated by a student as well as the number of low level functions articulated by a student. The maximum high or low level functions that a student could articulate was 4 for both Question 2 and 3. Cohen’s Kappa’s correlations between the scores given by the expert modelers follow: Fair agreement for Q2-High Level = 0.24, Q2-Low Level = 0.33, and Q3-High Level = 0.22, and Slight agreement for Q3-Low Level = 0.11. Again, rater scores were averaged together for analysis.

Third, Question 4 was scored by two graduate students at the same university using the rubric published by Nagel et al.28 An initial calibration occurred between the two graduate student scorers prior to each independently completing all model scoring. Calibration was comprised of both graduate student scorers scoring two models with an expert modeler and discussing each of the rubric items through the scoring process. A Pearson’s correlation of 0.91 was calculated, and again, scores between the two scorers was averaged. The highest score that a student could obtain from the rubric was an 18.

Table 2 provides the scores for student responses to Questions 1, 2, 3 and 4 for all four sections. The mean score and standard error for the scoring of the general correctness of each treatment group are plotted as Figure 4. Figure 5 provides the mean score and standard error for scoring of students ability to articulate high level and low level functions for Questions 2 and 3.

<table>
<thead>
<tr>
<th></th>
<th>Q1 Correctness</th>
<th>Q2 Correctness</th>
<th>Q2 High Level</th>
<th>Q2 Low Level</th>
<th>Q3 Correctness</th>
<th>Q3 High Level</th>
<th>Q3 Low Level</th>
<th>Q4 Rubric Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
<td>(standard deviation)</td>
</tr>
<tr>
<td>A</td>
<td>9.4 (1.4)</td>
<td>3.7 (0.6)</td>
<td>2.7 (0.9)</td>
<td>1.0 (0.9)</td>
<td>3.2 (1.1)</td>
<td>2.2 (0.9)</td>
<td>1.0 (0.8)</td>
<td>8.3 (3.8)</td>
</tr>
<tr>
<td>B</td>
<td>9.2 (1.3)</td>
<td>3.8 (0.4)</td>
<td>2.8 (0.7)</td>
<td>1.1 (0.8)</td>
<td>3.4 (0.8)</td>
<td>2.2 (0.8)</td>
<td>1.2 (0.9)</td>
<td>6.5 (2.7)</td>
</tr>
<tr>
<td>C</td>
<td>8.5 (1.2)</td>
<td>3.1 (1.2)</td>
<td>2.8 (1.1)</td>
<td>0.3 (0.4)</td>
<td>2.6 (1.2)</td>
<td>2.3 (1.0)</td>
<td>0.3 (0.5)</td>
<td>0.5 (1.5)</td>
</tr>
<tr>
<td>D</td>
<td>9.9 (0.3)</td>
<td>3.7 (0.7)</td>
<td>3.2 (0.8)</td>
<td>0.5 (0.5)</td>
<td>3.2 (0.7)</td>
<td>2.6 (0.9)</td>
<td>0.6 (0.9)</td>
<td>0.0 (0.0)</td>
</tr>
</tbody>
</table>
For Question 1, Kruskal-Wallis ANOVA is implemented since the data are not normally distributed and do not have homogeneity of variance \( (\chi^2=13.4, \text{df}=3, p=0.004) \). This demonstrates the sections are different. The control and experimental groups are not statistically different \( (\chi^2=1.49, \text{df}=1, p=0.22) \), and this is caused by Section C being lower than the experiment group and Section D averaging higher than the experimental group. For Question 2, a Kruskal-Wallis ANOVA \( (\chi^2=7.62, \text{df}=3, p=0.054) \) demonstrates that the sections are again different, and the control and experimental groups are different \( (\chi^2=4.036, \text{df}=1, p=0.045) \). The high and low level function results were not analyzed since the inter-rater agreement results are indicating that this metric is not reliable yet. For Question 3, a Kruskal-Wallis ANOVA demonstrates the sections are again different \( (\chi^2=7.30, \text{df}=3, p=0.063) \), and the control and experimental groups are also different \( (\chi^2=5.237, \text{df}=1, p=0.022) \). Question 4 is again not normally distributed and has unequal variances. For Question 4, the sections are different \( (\chi^2=54.6, \text{df}=3, p<0.01) \), and the control and experimental groups are different \( (\chi^2=52.8, \text{df}=1, p<0.01) \). Figure 4 and 5 provide plotted results of average scores for all scorings with error bars showing +/- one standard error.

**Figure 4.** Plotted results of average scores for general correctness of Questions 1, 2, 3, and 4 with error bars showing +/- one standard error.
5 Discussion and Future Work

The instruction variation, instruments, and accompanying data reported in this paper represents an initial step in determining how functional modeling instruction impacts students’ design abilities in a longitudinal setting. All students in this study learned that function is the enumeration of what a system is to do and formed as a verb-noun pair. Sections C and D only learned this meaning of function, while students in Sections A and B also received formal functional modeling instruction. Evaluation of the students’ understanding of function took the form of a four-question quiz (FunSkill). The quiz seeks to determine if students are able to 1) correctly distinguish functionality from other design attributes, 2) enumerate the functions of a well-known device, 3) enumerate the functions of an unknown device described by customer needs, and 4) generate a functional model.

Questions 1 through 3 of the FunSkill quiz were evaluated by engineering design professors while a pair of engineering graduate students evaluated Question 4 independently after training and calibration trials. The inter-rater agreement for Question 1 is 0.98, and the sections show statistical difference. But there was not a noted difference between the experimental group and the control group indicating that when taught about function as a concept, a student is able to discern a function statement from specifications, objectives, and other non-function statements. The evaluators agreed perfectly on the scores for 73 of the 76 samples, with the only variation occurring when it appeared a student had checked multiple boxes, scribbled in the selection boxes, or it was not readily clear which box they had selected. The high inter-rater agreement indicates that no further training protocols for assessing Question 1 are required. Also, the results from Question 1 support our hypothesis that the control and experimental group will equally be able to identify function statements.
Questions 2 and 3 are open-ended requiring students to describe a product functionally in the case of Question 2 or translate design objective statements into function statements in the case of Question 3. The scoring for both questions contained three parts. First, each was scored for general correctness by counting the number of correctly formed function statements listed by the student. Second, each was scored for the number of correctly formed “high-level” functions (i.e., broad, black box-level function statements describing an entire system-level function) listed by the student. Third, each was scored for the number of correctly formed “low-level” functions (i.e., functional model-level function statements describing a sub-system or component-level function). The inter-rater agreement for scoring general correctness of Questions 2 and 3 is high with a 0.81 and 0.85, respectively. While scoring these questions, the evaluators had some difficulty parsing what students had written; for example, some students would write a lengthy sentence containing which might be considered a function statement, or alternatively, a sentence would read more like a restatement of customer needs. Because of the open-endedness of the allowed responses, the evaluators find their agreement to be satisfactory. Statistical significance was also found between sections and between the experimental and control groups with Questions 2 and 3.

The coding for higher level and lower level functions in Questions 2 and 3 had very poor inter-rater agreement (less than 0.5), and as such cannot provide statistically meaningful results. To remedy the poor inter-rater agreement, the investigators plan to independently generate functional models of the products in question (finger nail clippers and dorm hand wash station) and then integrate the models into an agreeable representation of each system similar to the process used in Nagel et al.28 when calibrating for functional modeling scoring. Each function within the agreed upon models will then be discussed and classified as either high or low-level by the investigators such that they can be used as guiding examples for scoring future data sets. Consequently, the hypothesis related to Question 2 and 3 (The experimental group will perform better than the control group at enumerating sub-function-level function statements but no different at enumerating black box-level statements) can neither be rejected nor supported.

Question 4 was evaluated with the 18 point rubric as discussed in Section 4. The inter-rater agreement combined for Sections A, B, C, and D is 0.91 with statistical significance showing that students in Sections A and B far outperformed students in Sections C and D. In fact, a majority of students in Sections C and D did not even attempt to generate a functional model (scoring a 0), with some students simply drawing a picture of a bicycle. When correlation is calculated for only Sections A and B, agreement drops to a 0.71. Obviously there was no disagreement on scores for students in Sections C and D as they did not attempt to complete a functional model, thus the multitude of 0 scores help to raise the inter-rater agreement across all sections. In general, the models created by students in Sections A and B were sloppy and difficult to interpret, most likely due to time constraints of the quiz (~10 minutes total). The graduate students who scored the models reported having difficulty reading and understanding the student models; it is speculated that this difficulty may have contributed to the lower than desired inter-rater agreement. Of note also is that many students did not generate a black box model for Question 4; the lack of a black box model resulted in relatively low scores on the 18 point rubric as half of the questions are directly related to the black box model and its relationship to the functional model. Future studies will explore the differences in model quality when generated in low and high time-stress scenarios as well as at different times during the
semester to better understand how students learn and retain their functional modeling skills. The results of Question 4 support our hypothesis that the control group will be unable to generate black box or sub-functional functional models.

Overall, this data demonstrates that the FunSkills test is able to reliably measure the functional modeling skills of students. This paper also shows that students who receive more extensive training in functional modeling are able to differentiate functions from other design elements more effectively, enumerate functions for a device, and create functional models. The impact of these students’ functional modeling skills will be longitudinally studied as the students’ progress to other design classes and senior design. Students participating in this study will be tracked over the next two years to understand the impact of teaching function on engineering students’ design synthesis abilities. All students in this cohort will see other forms of qualitative modeling and system representation at later times through this study. This study is one of two studies occurring as a part of this project. Students at a large metropolitan public university, designated as High Doctoral Research by the Carnegie Foundation are also be participating. Studies at this second location are focusing on impact of teaching function on capstone design quality. Results of these studies are forthcoming.

Acknowledgements

This work is supported by the National Science Foundation through grants 1525449, 1525170, and 1525284. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of National Science Foundation.

References

27. Nagel RL, Bohm MR, Linsey JS. A Study on Teaching Functional Modeling in a Sophomore Engineering Design Course. 121st ASEE Annual Conference and Exposition; 2014; Indianapolis, IN.