Quantifying Differences Between Professional Expert Engineers and Engineering Students Designing: Empirical Foundations for Improved Engineering Education

Prof. Kurt Henry Becker, Utah State University

Kurt Becker is the current director for the Center for Engineering Education Research (CEER) which examines innovative and effective engineering education practices as well as classroom technologies that advance learning and teaching in engineering. He is also working on National Science Foundation (NSF) funded projects exploring engineering design thinking. His areas of research include engineering design thinking, adult learning cognition, engineering education professional development and technical training. He has extensive international experience working on technical training and engineering education projects funded by the Asian Development Bank, World Bank, and U.S. Department of Labor, USAID. Countries where he has worked include Armenia, Bangladesh, Bulgaria, China, Macedonia, Poland, Romania, and Thailand. In addition, he teaches undergraduate and graduate courses for the Department of Engineering Education at Utah State University.

Dr. John S. Gero, University of North Carolina, Charlotte

John Gero is Research Professor in Computer Science and Architecture at UNCC, Research Professor in Krasnow Institute for Advanced Study, and Research Professor in Computational Social Science at George Mason University. He was formerly Professor of Design Science, University of Sydney. He has edited/authored over 50 books and published over 650 research papers. He has been a professor of mechanical engineering, civil engineering, architecture, cognitive science, and computer science at MIT, UC-Berkeley, UCLA, Columbia and CMU in the USA, at Strathclyde and Loughborouh in the UK, at INSA-Lyon and Provence in France and at EPFL in Switzerland.

Dr. Morteza Pourmohamadi, Tabriz Islamic Art University

Morteza Pourmohamadi is an Assistant Professor in Industrial Design at Tabriz Islamic Art University. His research is mainly focused on cognitive studies of design and creativity.

Sarah Abdellahi, University of North Carolina, Charlotte
Lilian Maria de Souza Almeida M.Sc., Utah State University

Mr. Yuzhen Luo, Utah State University

Ph.D. student in Engineering Education at Utah State University.
Quantifying Differences Between Professional Expert Engineers and Engineering Students Designing: Empirical Foundations for Improved Engineering Education

Introduction

One of the primary goals of engineering design education is to equip students with the capability of becoming expert design engineers. To develop this capability in students, educators require a detailed knowledge of the cognitive behavior of both undergraduate students and expert design engineers. However, the lack of information about the cognitive behavior of expert design engineers led to a gap between competencies developed in universities and those needed to become an expert in the field; the process to quantify and verify how students acquire “the ability to do expert work” is inadequately studied in engineering.

The purpose of this research is to begin to characterize engineering learning so that we may begin to identify potentially novel pathways to approach the cognitive transformation from novice to expert in engineering education. This project measures and compares the design thinking of dyads of freshmen engineering students, dyads of senior engineering students, and dyads of professional expert engineers through a study of their cognitive processes while designing. It uses tools and processes developed in previously funded NSF projects to provide a uniform basis for comparing students and professional experts that is independent of the educational and experiential background of the participants.

Outcomes of this research provide a cognitive foundation to inform and improve engineering education models while expanding our understanding of how students evolve to acquire expert-level design skills. The results inform leaders in engineering education and developers of instructional materials and curricula, as well as teachers and designers planning classroom strategies, of initiatives in formal engineering education.

Three methodologies are used to characterize and model the effects of education and experience on engineering students’ and expert designers’ design cognition. The methodologies are drawn from: 1) design theory: design ontologies, 2) cognitive science: protocol analysis and cognitive style, and 3) statistical modeling: standard statistical analysis, Markov modeling, and problem-solution index. Given that different designers with varying education and experience backgrounds, and designing for a variety of requirements under different conditions, the Function-Behavior-Structure (FBS) ontology methodology from design science is utilized as a means of characterizing designing in a uniform way that is independent of the designer, the design task, and the design situation (Kan and Gero, 2017). Consequently, the FBS ontology enables the development of more complete models to articulate and quantify the differences between the cognitive behaviors of novice and expert engineering designers.

![Figure 1. The FBS ontology](image-url)
The FBS Ontology

The FBS ontology of designing has been used in multiple disciplines and one that transcends individual designers, the design task, the design environment, and whether designing individually or in teams (Branki, 1995; Hofmeister, et al., 2007; Jiang, 2012; Kruchten, 2005; Robin, et al, 2007; Van Wie, et al., 2005; Visser, 2006).

The FBS ontology (Gero, 1990; Gero & Kannengiesser, 2014) models designing in terms of three classes of ontological variables: function, behavior, and structure plus a design description, Figure 1 shows the FBS ontology. The goal of designing is to transform a set of functions, driven by the client requirements (R), into a set of design descriptions (D). The function (F) of a designed object is defined as its intended purpose or teleology; the behavior (B) of that object is either derived (Bs) or expected (Be) from the structure, where structure (S) represents the components of an object and their relationships.

Designers decide which behaviors (B) are significant and needed to assess the designs they produce. So, B can be subdivided into two sub-categories: the behaviors the designer expects the design to have (Be) and those that are measured from the design (S) itself and called behavior from structure (Bs). Different functions for the same design produce different expected behaviors that generate different structures. An example of two different functions invoking different behaviors and different structures for the same design using a cell phone is shown in Figure 2.

![Figure 2](image)

Figure 2. An example of functions (F), expected behaviors (Be) and structures (S)

The FBS coding scheme can be summarized, using the design terminology embodied in Figure 1. This produces six codes for the design issues (segments) and those six codes can be combined to produce eight design processes, Tables 1 and 2.

**Design Cognition Research**

Much of engineering education research in design is dominated by explorations of design teaching, although recently there have been cognitive studies of designers that have been aimed
at elucidating design-thinking behavior. These studies have fallen into five methodological
categories: questionnaires, interviews (Cross & Cross, 1998); input-output experiments, where
the designer is treated as a black box which produces the behaviors in the outputs for changes in
inputs (Purcell, Williams, et al, 1993), anthropological studies (Lopez-Mesa & Thompson, 2006),
and protocol studies. While each of these methods produced interesting results, the most
promising method is protocol studies. It has become the basis of the current cognitive study of
designers (Atman, et al. 2008; Badke-Schaub et al 2007; Becker & Mentzer, 2012; Christensen &
Schunn 2007; Gericke, et al 2007; Gero, Kan & Jiang, 2014; Kavakli & Gero, 2002; McDonnell
& Lloyd, 2007; McNeill, et al, 1998; Song, 2014; Suwa, et al, 1998; Suwa, Gero & Purcell,

Table 1. FBS Codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Design Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Formulation</td>
</tr>
<tr>
<td>F</td>
<td>Synthesis</td>
</tr>
<tr>
<td>Bs</td>
<td>Analysis</td>
</tr>
<tr>
<td>Be</td>
<td>Documentation</td>
</tr>
<tr>
<td>S</td>
<td>Evaluation</td>
</tr>
<tr>
<td>D</td>
<td>Reformulation I</td>
</tr>
</tbody>
</table>

Table 2. FBS Processes

<table>
<thead>
<tr>
<th>Design Process</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation</td>
<td>R&gt;F, F&gt;Be</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Be&gt;S</td>
</tr>
<tr>
<td>Analysis</td>
<td>S&gt;Bs</td>
</tr>
<tr>
<td>Documentation</td>
<td>S&gt;D</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Be&lt;&gt;Bs</td>
</tr>
<tr>
<td>Reformulation I</td>
<td>S&gt;S</td>
</tr>
<tr>
<td>Reformulation II</td>
<td>S&gt;Be</td>
</tr>
<tr>
<td>Reformulation III</td>
<td>S&gt;F</td>
</tr>
</tbody>
</table>

This project builds on previous NSF-funded projects that looked at the longitudinal development
of design cognition of undergraduate engineering students across two contiguous years
(Williams, Lee, Gero & Paretti, 2013). Results from a pilot study at Utah State University show
there is a significant gap between the cognitive behavior of novice and professional expert
engineering designers (Song, 2014). This project makes use of those results and focuses on gaps.
It brings together the beginning of engineering education (freshmen), works with students
completing engineering education (seniors), and completes the longitudinal development of
engineering design by studying professional experts.

Expert engineers are those who have at least 10 years and 10,000 hours of professional
experience (Cross, 2004; Dufresne, Gerace, et al, 1992; Ericsson, Charness, Feltovich, &
Hoffman, 2006; Kaufman, & Kaufman, 2007; Kavakli, & Gero, 2002). The design cognition of
expert design engineers is inadequately characterized. The focus of previous studies on expert
design behavior have been on case studies that produced qualitative results (Ahmed 2001;
design cognition studies indicated that there are commonalities as well as differences between
students and professional designers (Gero, Kannengiesser & Pourmohamadi, 2012). Understanding difference between novices as developing learners and expert target performance is essential to identify appropriate learning experiences to reduce this performance gap.
Research Design

Figure 3 shows the research design including inputs, the process, analysis and outputs. Using the function–behavior–structure (FBS) ontology, empirical research is collecting data from the verbal protocols of 60 undergraduate engineering students in teams of two, and 20 professional expert engineers in teams of two, while designing. A cohort of 20 is sufficient to provide a statistically reliable dataset to measure any differences when the differences are based on a single variable. Expert engineers are selected from a design companies in Seattle, Washington, Los Angeles, California and Salt Lake City, Utah. Having the expert team members come from across the country gives a representative cross-section of the US.

Figure 3. Research design showing inputs, process, analysis and outputs

Design Task

In this research, all teams completed the same functional level engineering design task. The task is related to a window design that is not familiar to either the students or the experts. Under a one-hour time constraint, teams were encouraged to use brainstorming and present a final sketch of their design upon completion. The teams had access to additional resources including the Americans with Disabilities Act (ADA), ADA Accessibility Guidelines for Buildings and Facilities, a video of the window, and a detailed description of the window’s parameters to support their design process.

Data Collection

Thirteen teams of freshmen, eleven teams of seniors and thirteen teams of professionals form the source data. The teams of professionals had a larger diversity of background knowledge domains than was originally intended because of the greater difficulty in recruiting professionals than expected. All participants were volunteers who were given a gift card for their time. Individual teams were videoed as they designed. Since they are collaborating, the team members naturally verbalize. The verbalizations and gestures are transcribed and the resulting transcripts were coded by two independent coders using a 6 element coding scheme based on the FBS ontology: R, F, Be, S, Bs and D. The coders then arbitrated between themselves where there are disagreements in their codings. In those instances where the coders could not come to an arbitrated agreement, the project’s principal investigator carried out the final arbitration in conjunction with the session’s coders. Four coders were used and were rotated between the
coding of the design session to reduce coder bias when working with the same coder continuously. Rather than using Cohen’s kappa as the measure of inter-coder reliability, this project measures each coder’s agreement with the final arbitrated code. The average agreement across all design sessions between the coders’ codings and the final arbitrated codings was 86.8%.

The effect of this activity was to turn the video of each design session into a sequence of codes, where each code is associated with a semantic designing concept (Kan & Gero, 2017). These sequences of codes became the data for later analyses.

**Statistical Modeling**

Multiple classes of statistical analysis techniques are employed to obtain models from the data sets of the final protocol. Based on the information produced by these techniques, quantitative comparisons between the different levels of design experience can be made. The results of using two classes of statistical modelling are presented here. The results from other modelling approaches will be presented in future papers.

*Correspondence Analysis*: Correspondence analysis (Husso, Le & Pages, 2011) is a form of principal component analysis applied to categories rather than individuals. It produces a semi-qualitative representation of the relationships between the categories from which qualitative assessments can be made.

*Standard Statistical Analysis*: This generates the statistical distributions along with their variances of codes in segments in each of the final protocols This provides the foundation for the characterization of the design cognition of participants.

*Specialized Analyses*: Two additional measures, cumulative occurrence models and P-S index, have been developed (Kan & Gero, 2017) and are used here.

*Cumulative occurrence models*: The cumulative occurrence $c$ of design issue $x$ at design step $n$ is defined as $c = \sum_{i=1}^{n} x_i$ where $x_i$ equals 1 if design step $i$ is coded as $x$ and 0 if design step $i$ is not coded as $x$. Plotting the results of this equation on a graph with the segments $n$ on the horizontal axis and the cumulative occurrence $c$ on the vertical axis will show the occurrence of the design issues (Kannengiesser, et al., 2013; Pourmohamadi, 2010).

The $P-S$ index is a meta-level model generated by dividing the designing activities into two cognitive spaces: problem space and solution space. The Problem-Solution index or P-S index is a quantitative measure of the cognitive effort distributed between these two spaces (Jiang, et al., 2012). And is calculated using Equation (1). P-S indexes with a single value facilitate comparisons across multiple sessions and across sessions involving different situations.

$$P-S \text{ index (design issue)} = \frac{\sum(\text{Problem-related issues})}{\sum(\text{Solution-related issues})} = \frac{\sum(R,F,B, e)}{\sum(Bs,S)}$$

(1)

The mappings of the design issues and design processes into the problem space and solution space are given in Table 3.
Table 3. Mapping Design Issues and Design Processes onto Problem and Solution Spaces

<table>
<thead>
<tr>
<th>Problem/Solution Space</th>
<th>Design Issue</th>
<th>Design Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning about Problem</td>
<td>Requirement (R)</td>
<td>1 Formulation</td>
</tr>
<tr>
<td></td>
<td>Function (F)</td>
<td>8 Reformulation II</td>
</tr>
<tr>
<td></td>
<td>Expected Behavior (Be)</td>
<td>7 Reformulation III</td>
</tr>
<tr>
<td>Reasoning about Solution</td>
<td>Behavior from Structure (Bs)</td>
<td>2 Synthesis</td>
</tr>
<tr>
<td></td>
<td>Structure (S)</td>
<td>3 Analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 Evaluation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 Reformulation I</td>
</tr>
</tbody>
</table>

A publicly available software tool that carries out the standard statistical analysis and other modeling on protocol data called LINKODER (http://www.linkoder.com/) (Gero, Kan & Pourmohamadi, 2011) was used.

Results

Correspondence Analysis

In order to determine if there are categorical differences between the design cognition of the freshmen, seniors and professionals, we carried out a correspondence analysis on all the data for each of these three cohorts, Figure 4. The six cognitive codes are treated as categories and the results of their correspondence analysis are also shown in Figure 4. This allows for the determination of whether the codes are indicative of distinct categories of cognitive concepts. The two dimensions generated by the correspondence analysis respectively cover 74.9% and 25.1%, i.e., a total of 100% of the variance implying that these two dimensions are sufficient to describe the categories. In correspondence analysis, which is a form of dimensional reduction, the dimensions are in mathematical space not in the original problem space. The dimensions are those that minimize the variance of the data coverage and are used to assess qualitative differences between the categories.

We start by looking at the locations of the codes in these two dimensions. We note that R (requirements) and F (functions) sit in the same quadrant implying that categorically they are close to each other. This matches our understanding of their roles in designing. R, F and Be (expected behavior) all sit on the positive side of Dimension 1 implying that they are categorically distinct from the other three codes. This matches our understanding of their roles in designing in that they are all associated with the problem while the other three are associated with the solution. Bs (behavior from structure) sits in a different quadrant to S (structure) and D (documents). This also matches our understanding of their roles. Structure is the thing and Bs is derived from S and is dimensionally different. This categorization of the codes used to divide and structure each design session provides an empirical foundation for their use in this research.
Figure 4. The results of the correspondence analysis of the all the data for the freshmen, seniors and professionals. The six cognitive codes are treated as categories and their results are also shown.

Turning now to the freshmen, seniors and professionals, we note that they sit in three separate quadrants in Figure 4 implying that they are categorically different. The professionals and seniors sit in the positive side of Dimension 1, implying that in that dimension they have some similarities. The freshmen sit on the negative side of Dimension 1, implying that in that dimension they are different to both the seniors and the professionals. The professionals and seniors sit on the opposite sides of Dimension 2 implying that they are different, whilst the freshmen sit in between.

These results establish that there are categorical differences between the design cognition of freshmen, seniors and professionals. We now turn to descriptive statistics to examine whether these differences can be isolated.

**Standard Statistical Analysis**

Here we will present statistical results at the aggregate level where we treat design sessions as a single unit. In a later paper we will examine the design sessions at a smaller level of granularity and will include time across the sessions.

The distributions of the cognitive design issues for the three cohorts are shown in Figure 5.
All three cohorts spent three-quarters of their cognitive effort on the design issues of structure and behavior from structure. The design issue of structure dominated the cognitive effort of all three cohorts. This is similar to results from previous studies of designers (Gero, Kannengiesser and Pourmohamadi, 2014). Later in the paper we present the results of testing whether there are statistically significant differences in the distributions of the design issues between the cohorts.

The distributions of the cognitive design processes for the three cohorts are shown in Figure 6.
**Figure 6.** Distributions of the cognitive design processes, expressed as percentages of all design processes for freshmen, seniors and professionals.

For all three cohorts the dominant process is Reformulation 1 – operating in the structure space. This is followed by Analysis for all three cohorts. Later in the paper we present the results of testing whether there are statistically significant differences in the distributions of the design processes between the cohorts.

We then examined the rate at which participants expended cognitive effort on the design session. This is measured by the cumulative occurrence of design issues (Kan and Gero, 2017).

**Figure 7.** Mean slopes of cumulative issues of freshmen, seniors and professionals.

ANOVA tests ($\alpha = 0.05$) were carried out for the cognitive design issues distributions, design process distributions, cumulative design issues and quartile P-S indexes. These are presented with their $p$-values in Tables 4, 5, 6, and 7 respectively. The Between Group degrees of freedom is 2 and Within Group degrees of freedom is 31. Items marked with an asterisk (*) indicate that there may be insufficient data to justify their $p$-values.

The results of testing for differences in design issue distributions are shown in Table 4.

**Table 4.** The ANOVA $p$-values for Issue Distributions

<table>
<thead>
<tr>
<th>Issues Distribution</th>
<th>$p$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be</td>
<td>0.0007</td>
</tr>
<tr>
<td>Bs</td>
<td>0.0005</td>
</tr>
<tr>
<td>S</td>
<td>0.1147</td>
</tr>
<tr>
<td>F*</td>
<td>0.6061</td>
</tr>
<tr>
<td>R*</td>
<td>0.0028</td>
</tr>
<tr>
<td>D*</td>
<td>0.9081</td>
</tr>
</tbody>
</table>
The results in Table 4 indicate that there are statistically significant differences between the three cohorts in the distributions of the design issues of Be and Bs.

The results of testing for differences in design process distributions are shown in Table 5.

**Table 5. The ANOVA p-values for Design Process Distributions**

<table>
<thead>
<tr>
<th>Process Distribution</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation</td>
<td>0.0136</td>
</tr>
<tr>
<td>Synthesis</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Analysis</td>
<td>0.0568</td>
</tr>
<tr>
<td>Evaluation</td>
<td>0.0607</td>
</tr>
<tr>
<td>Documentation</td>
<td>0.5277</td>
</tr>
<tr>
<td>Reformulation 1</td>
<td>0.0628</td>
</tr>
<tr>
<td>Reformulation 2</td>
<td>0.0004</td>
</tr>
<tr>
<td>Reformulation 3*</td>
<td>0.3374</td>
</tr>
</tbody>
</table>

The results shown in Table 5 indicate that there are statistically significant differences between the three cohorts in the distributions of the design processes of Formulation, Synthesis and Reformulation 2, along with weak differences in the design processes of Evaluation and Reformulation 1.

The cumulative design issues are a measure of the time-distribution of cognitive effort across a design session as compared to just the distributions as shown in Figure 5, which have no time dimension. The results of testing for differences in the slopes of the cumulative design issues graphs are shown in Table 6.

**Table 6. The ANOVA p-values for Cumulative Design Issues**

<table>
<thead>
<tr>
<th>Cumulative Design Issue</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be</td>
<td>0.0006</td>
</tr>
<tr>
<td>Bs</td>
<td>0.0013</td>
</tr>
<tr>
<td>S</td>
<td>0.1723</td>
</tr>
</tbody>
</table>

The results in Table 6 indicate that there are statistically significant differences between the three cohorts in the cumulative design issues of Be and Bs.

The sequential P-S indexes across different sections of a designing session generate a time-based “signature” of the cognitive style of the activity. When the session is divided into quartiles, the P-S index for each quartile is calculated, and used in a sequence of temporally ordered P-S indexes to represent the design style changes during the session. The results of testing for differences in the P-S index across quartiles are shown in Table 7.
Table 7. The ANOVA p-values for Quartile P-S Indexes

<table>
<thead>
<tr>
<th>Quartile P-S Index</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Quartile</td>
<td>0.0090</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.0024</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.0659</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>0.0289</td>
</tr>
</tbody>
</table>

The results in Table 7 indicate that there are statistically significant differences in the 1st, 2nd and 4th quartiles of the P-S index.

Conclusions

Protocol studies were carried out of three cohorts: freshmen, seniors and professionals designing with the same design prompt. Results indicate that there are categorical differences between each of the cohorts. Preliminary statistical testing for differences indicated statistically significant differences in design issue distributions and design process distributions between the cohorts while designing. Since we found categorical differences we would expect to find statistically significant differences. We then calculated the cumulative occurrences of the design issues and found statistically significant differences between the rates of cognitive effort of students and professionals. We intend to carry out further statistical modeling such as Markov modeling (below). The professional cohort represented a more diverse group of expert designers than the original experiment called for due to the availability of the professionals and any effect of this wider diversity will be investigated further.

Since Markov modeling is not commonly used, we present its application to design protocols. Design styles and designer’s strategies in terms of repeated processes can be assessed by building Markov models of the transitions between design issues and design processes (Kan & Gero, 2017). Markov models (Kemeny & Snell, 1960) generate the probability of a particular design issue following another particular design issue. Markov models to represent cognitive design style have been used across multiple domains (Kan & Gero, 2017; Jiang, 2012; Pourmohamadi, 2013) and is one foundation for measuring quantitative differences between students and experts. Richer design patterns can be found using second- and third-order Markov analysis.

These results need to be related to specific curricula to determine whether and where curriculum changes could be made to aid students to improve their design cognition in the course of forming themselves into engineers.

The results from these types of empirical investigations inform leaders in engineering education and developers of instructional materials and curricula, as well as teachers and designers planning classroom strategies, of initiatives in formal engineering education. The development of educational strategies are explored and developed through a workshop of engineering design educators to move students along a trajectory towards expert design behavior.
Acknowledgement
This material is based upon work supported by the National Science Foundation under Grant Nos. 1463873 and 1463809. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.”

References


