Advances in Engineering Education

SUMMER 2015

An Engineering Educator's Decision Support Tool for Improving Innovation in Student Design Projects

NUR OZGE OZALTIN

MARY BESTERFIELD-SACRE

AND

RENEE M. CLARK
University of Pittsburgh
Pittsburgh, PA

ABSTRACT

Learning how to design innovatively is a critical process skill for undergraduate engineers in the 21st century. To this end, our paper discusses the development and validation of a Bayesian network decision support tool that can be used by engineering educators to make recommendations that positively impact the innovativeness of product designs. Our Bayesian network model is based on Dym's design process framework and actual design process data collected from 26 undergraduate engineering capstone teams over multiple terms. Cross validation using all available outcomes data and a sensitivity analysis showed our model to be both accurate and robust. Our model, which is based on data from teams that produced both innovative and non-innovative products, can be used to formatively assess the process used by a design team and the level of innovativeness, thereby contributing to more innovative final design outcomes.

Key Words: decision support tool, Bayesian network model, engineering design education

INTRODUCTION

To capture and retain market share in the modern business environment, today's organizations must meet or exceed customer expectations through product innovation (Shen et al., 2000). Innovation includes the introduction of new materials, components, and/or manufacturing processes; design changes employed to reduce manufacturing costs; and new applications of existing technology (Schumpeter, 1934).

Solid knowledge of innovative design is critical for successful, professional contribution within the workforce; therefore, participation in student design projects such as senior capstone experiences is an integral part of the undergraduate engineering curriculum. These activities are invaluable learning experiences and equip future engineering professionals with the ability to design new products and services with innovative features in rapidly changing and highly competitive markets.

Thus, instructors who teach innovative design techniques would benefit from a decision support tool for guiding teams to carry out those activities and processes that best contribute to the production of innovative products. To this end, our paper describes the development of such a tool that allows engineering design instructors to 1) predict the innovativeness of a design artifact given the team's activities as well as 2) provide guidance on activities that would contribute to a more innovative design outcome. This tool, implemented as a Bayesian network model, was developed based on empirical process-level activity data collected from undergraduate teams that produced both innovative and non-innovative capstone design products. Specifically, the activities of 26 senior bioengineering design teams were journaled using a secured, online survey system over a 23 to 24 week period as they progressed from an initial design concept to a working prototype.

Subsequently, using this process-level activity data and Dym's design process framework, three separate Bayesian network models for the early, middle, and late phases of the design process were built using the GeNie decision modeling software. Bayesian networks allow for determining the probability that an upstream event or activity occurred given that a downstream event, such as an innovative outcome on a capstone design project, also occurred (Genie Documentation, 2014). Thus, with a Bayesian network and ultimately Bayes Theorem, on which this network is based, we can determine the probability that a certain design process activity was performed to a certain degree in a certain project phase given that the design was ultimately rated as "innovative" or "non-innovative." This type of analysis can lead to data-driven recommendations from instructors to students. Our work began as a doctoral dissertation, with subsequent refinement and analysis by the additional authors (Ozaltin, 2012).

LITERATURE REVIEW AND DESIGN FRAMEWORK

Engineering Design and Design Frameworks

Design is a central and distinguishing engineering activity (Simon, 1996). It is a complex process versus a single isolated action that has a collectivist nature (Okudan & Mohammed, 2006). There is also no universally agreed-upon definition of design (Hyman, 2003). The literature has many simplified step-by-step models and frameworks of the design process based on the ABET definition, which presents the aspects of an ideal engineering design process (ABET Criteria, 2013). Hyman breaks the design process into the following nine steps: recognizing the need, defining the problem, planning the project, gathering information, conceptualizing alternative approaches, evaluating the alternatives, selecting the preferred alternative, communicating the design, and implementing the preferred design (Hyman, 2003). Another framework proposed by Pugh is based on a design core including the market, specification, concept design, detailed design, manufacture, and marketing (Pugh, 1990). Atman et al. compare freshman and senior engineering design processes and describe the design steps as follows: identification of the need, problem definition, gathering of information, generation of ideas, modeling, feasibility analysis, evaluation, decision, and communication (Atman et al., 1999). Dominick et al. define engineering design as an iterative process, and they segment it into the following four main phases: defining the problem, formulating solutions, developing models and prototypes, and presenting and implementing the design (Dominick et al., 2001). Several other engineering design models and frameworks can be found in the literature (Pahl et al., 2007; Lewis & Samuel, 1989; French, 2010; Cross, 2001), although none of these models or frameworks is universally accepted by the engineering community (Hyman, 2003). We chose Dym's engineering design framework, shown in Figure 1, as our theoretical model to generalize and thereby simplify

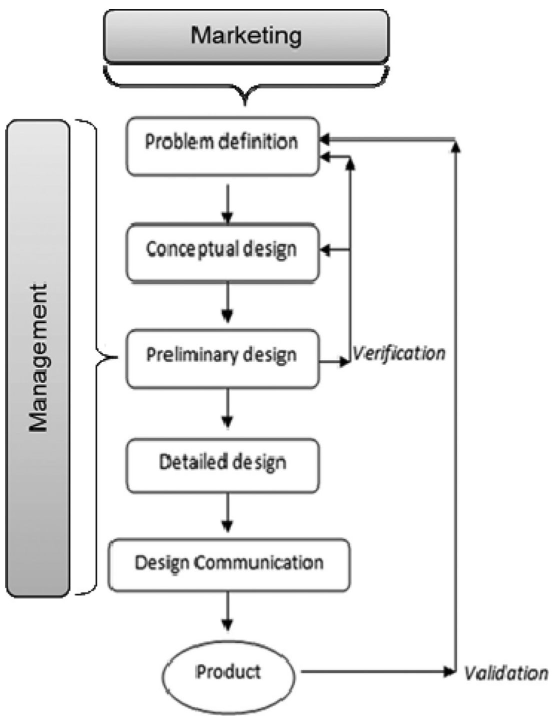

Figure 1. Feedback and Iteration in Dym's Design Process.

the large number of design activities used by our students (Dym & Little, 2004). We selected this model for two reasons. First, Dym et al.'s definition of engineering design matched the goals of the senior capstone project well. Their definition is as follows: "*Engineering design is a systematic, intelligent process in which designers generate, evaluate and specify concepts for devices, systems, or processes whose form and function achieve clients' objectives or users' needs while satisfying a specified set of constraints"* (Dym & Brown, 2012, p. 16; Dym et al., 2005, p. 104). In addition, Dym's framework is flexible enough to be applied in different fields of engineering, including systems engineering, but still detailed enough to model important design activities, including iterations. Dym's model was also preferable to other design process models in the literature because it aligned with both the engineering design and engineering education disciplines and was well-suited to the data collected. Since our research was focused on both design and product realization, we expanded Dym's model by adding the *marketing* and *management* categories, since many product realization projects incorporate these activities. Hence, our overall process was described by eight categories (i.e., the original six categories in Dym's design process model and two product realization categories).

For this research, we adapted 89 design activities for collecting design process data throughout an academic year (Golish et al., 2008). These activities can be organized into engineering design stages based on the literature, such as opportunity identification, design and development, testing and preproduction, and introduction and production. In studies of design, it is common to generalize design activities into a smaller set of categories and/or cognitive operations. For example, the categories of exploration, generation, comparison, and selection have been used (Stempfle & Badke-Schaub, 2002). The challenge with this approach is its dependency on the information about the design activities and the fact that these activities often occur in cycles or iterations (Ha & Porteus, 1995; Krishnan et al., 1997).

Bayesian Networks and Their Applications

Within the broad area of engineering design, our research focused specifically on creating a tool to influence innovative design outcomes. We ultimately used a Bayesian network (BN) to develop a decision support tool to increase the likelihood of innovative outcomes in design settings. A Bayesian network is a probabilistic graphical model that represents a set of variables as circular nodes and their conditional dependencies or interactions as arcs or arrows. A BN allows for forward and backward inference under uncertainty given known evidence and is useful for analyzing "what-if" scenarios, even those that are not observed in practice (Jensen & Nielsen, 2007; Yannou et al., 2013; Genie Documentation, 2014). We used the GeNie software to create our Bayesian network model (Genie Documentation, 2014). This software provides a development environment for building graphical decision models. Although Bayesian networks are formulated using only chance nodes, the "set evidence" property of GeNie allows a chance node to be treated as a "decision node" by setting the evidence to a chosen state.

Although the most popular application area for Bayesian networks is medical decision making, especially verification of a diagnosis, they have a wide range of applications in finance (e.g., market analysis), reliability (e.g., processor fault diagnosis), and defense (e.g., automatic target recognition). Bayesian networks have also been applied to engineering design problems, including improvement of the early design stage by addressing uncertainties in component characteristics and compatibility (Moullec et al., 2013). This model also contributed to innovation by ensuring product feasibility and reducing the design risk. They focused on the early design stage (i.e., conceptual design) and determined the probabilities based on expert opinion. In contrast, our BN model encompasses the stages of conceptual design through prototype development and was built using actual design-team data to estimate the probabilities.

Another application of BN's to engineering design was an evaluation of innovation by considering industrial contexts (Yannou et al., 2013). These authors performed an empirical study to identify the factors related to design and analyzed the influence of these factors on the quality of the problem setting and subsequently the problem-solving process as well as the quality of the innovative project outcome. In comparison, our model offers suggestions for the utilization levels of the design activities that may lead to more innovative design outcomes.

METHODS

Data Collection

The data used for developing our decision support tool was collected from bioengineering senior capstone design teams during the 2007-08 and 2008-09 academic years. Twenty-six teams participated, with 18 teams from an engineering school in the Mid-Atlantic region and eight teams from an engineering school in the Midwest. The design projects were similar in nature, in which all students had to design a biomedical product or device. Where possible, we minimized variability between the two institutions. The number of students per team varied from three to five, and the students were paid for their participation in the study. Students were surveyed twice per week through a secure online system to collect quantifiable data about their design activities. Within each of four design stages, a student could select up to three activities he/she had worked on. This number was arbitrarily set but was believed to be sufficient given the three to four day interval between the surveys. If the student had not worked on the project since the last survey, he/she could select

"I have not worked on the design." Each student completed the survey up to a total of either 45 or 48 times, depending on his/her school. As discussed, the entire set of activities was based upon the work of Golish et al. and was further refined by the capstone instructors (Golish et al., 2008). Students were trained in the meaning of the activities and were provided with a definition sheet for easy reference.

It was assumed that students selected the activities and answered the open-ended questions in good faith. It is our belief for multiple reasons that students were honest in providing data. During their initial training session, students were informed that their answers would not be shared with the instructors and would not affect their grades. We utilized the training session and the assurances we gave to students at the time as a means to alleviate the Hawthorne effect (McBride, 2010). Students had the option to select "I have not worked on the design," which was chosen 129 times during the project timeframe. Further, while reviewing the data, students appeared to be selecting logical activities and writing detailed reflections. Their responses did not appear to be cursory in any manner.

Although each design was graded according to the instructors' course criteria, each instructor also evaluated the projects using a common rating scale consisting of five criteria, including one for assessing the innovativeness of the final product. The five criteria used to assess the design projects were technical performance and standards, documentation, innovation, working prototype, and overall impact on the market or to the client. Each criterion also contained sub-criteria derived from the literature that the instructors collectively agreed upon. The innovation sub-criteria were based upon Schumpeter's definition of innovation; therefore, the innovativeness of the products was evaluated based on 1) new applications of existing technology, 2) use of new materials or components, 3) introduction of new manufacturing processes, and 4) design changes that reduce costs.

The scoring rubric was derived from the VentureWell (formerly National Collegiate Inventors and Innovators Alliance) BMEIdea Competition (BMEidea, 2014). Using this as a starting point, the researchers and instructors of the bioengineering capstone courses iteratively revised the rating scale to arrive at an agreed-upon set of defined attributes. Unfortunately, each instructor rated his/her teams' products and reports separately, and this is a limitation of our work. In terms of qualifications, three of the instructors were biomedical engineering faculty, and one instructor was a bioengineering faculty member as well as a co-founder of an innovative medical products company. This faculty member also has over ten medical-related patents. In addition to these qualifications for assessing the innovativeness of bioengineering projects, the literature supports the validity and consistency of faculty ratings on student and team performance when developed in an iterative manner (Stiggins, 1999; Moskal & Leydens, 2000; Callahan et al., 2000).

To determine whether our results were reflective of "innovation" versus what might constitute "good design," the correlation of innovation with each of the other criteria was calculated. With the documentation criterion, there was little variation, as all teams had similar scores; thus, documentation was not considered. The correlations of innovation with the other three criteria were between 0.42 and 0.52; thus, innovation was at best moderately correlated with the other components of design. These three correlations were significantly different from zero at α=0.05 (Ozaltin et al., 2015).

The rubric scores ranged from "1" (poor) to "5" (excellent). For this research, products having a score of "4" or "5" on the innovation criteria were considered innovative; and conversely, products having scores of "1" or "2" were considered non-innovative. Overall, there were eight innovative and eight non-innovative products from the 26 teams, with 10 products that were rated as neither innovative nor non-innovative.

Model Development

The 89 possible activities used by the students were classified into the eight categories of our theoretical model by an experienced research team consisting of five individuals in the field of design and product realization. The research team members first individually and then collectively arranged all activities according to Dym's model. Discrepancies between members were discussed to determine the best fit of the activities to the categories. In some cases, it was determined that certain activities could be placed in multiple categories.

The element of time was an important aspect of this research. To obtain stronger and more specific results, the project timeline was divided into three phases - early, middle, and late. Given the relatively long time frame of the students' design process (i.e., multiple terms), three separate BN models representing the early, middle and late phases were created. As seen in Figure 2, a five-day transition period was used between the phases to prevent rigid borders. A partial membership rule was applied for those activities observed in the transition period (Ozaltin et al., 2015). The early and late phases had four time epochs each, and the middle phase contained five. A time epoch was representative of approximately two to two and a half weeks of design activity.

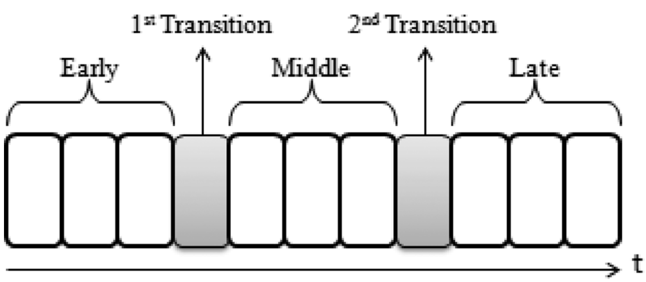

Figure 2. Timeline of the Design Process.

We considered several modeling approaches for developing our tool. A decision-based approach was required, as engineering design teams continuously make decisions throughout the process (Lewis et al., 2006). The model had to consider any history, since past design activities are critical in determining current and future activities. Also, although activity selection and usage influence innovativeness, they do not guarantee an innovative output. Thus, the proposed model had to allow for uncertainty. Lastly, given the data collected, there were no intermediate rewards but only a final reward (i.e., the final prototype score). Markov chains (MC), Markov decision processes (MDP), and influence diagrams (ID) were each evaluated as candidate modeling techniques. Influence diagrams supported all of the requirements, including a decision-based approach, maintenance of history, consideration of uncertainty, and accommodation of an end reward. In particular, a Bayesian network, a special case of an influence diagram, was chosen as the modeling technique.

Our Bayesian network depicts how design teams used Dym's design process across time to achieve a design artifact, albeit innovative or non-innovative. Using Dym's extended model, our BN contains eight variables, as shown in Figure 3, where each variable, or design category, is represented by a node in the model. Each node has three states representing the level of usage by a design team: *low, medium,* or *high*. The variables are repeated across the time epochs *t*.

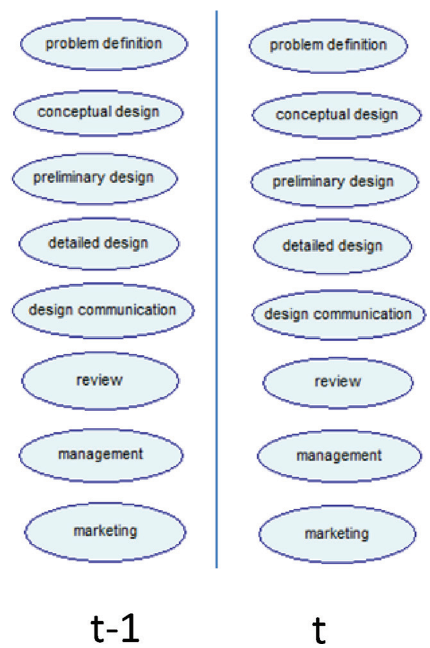

Figure 3. Bayesian Network Variables.

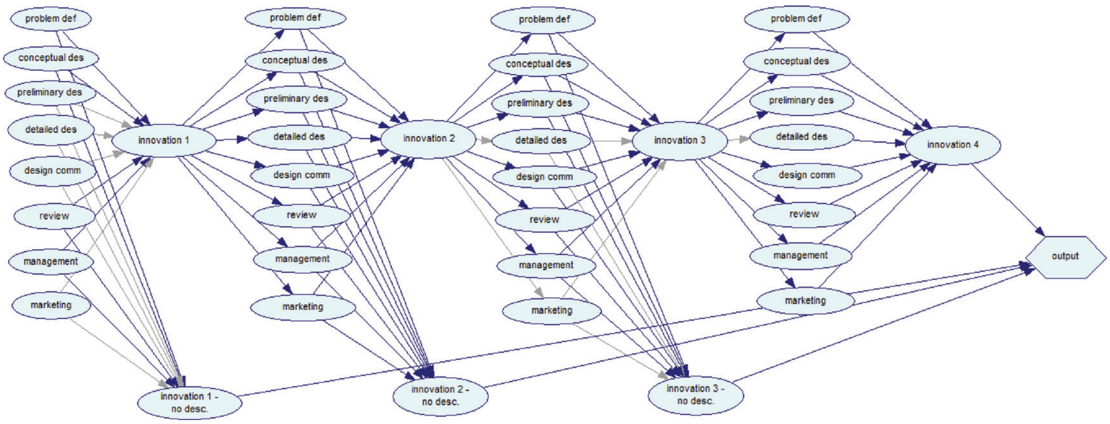

Figure 4. Bayesian network model.

Figure 4 shows the various nodes of the early phase model, which appear in all four time epochs. The final output of the model is a node consisting of two states: innovative vs. non-innovative. Given the large number of nodes in the model, there were a very large number of potential relationships or dependencies (i.e., arcs) among the nodes. We determined that the majority of all possible relationships, as quantified through conditional probabilities, did not occur and were therefore not represented in our data. In addition, when all possible relationships were represented, the model became overly complicated and unwieldy. Therefore, to simplify as well as enhance the model, we alternatively added innovation inter-nodes with two states after each time epoch to efficiently connect the epochs, as shown in Figure 4.

Unfortunately, when evidence is entered into a subsequent epoch, the effect of the previous time epochs is virtually lost. This meant, for example, that prior to the addition of the inter-nodes, evidence set in the first and fourth time epochs did not have the same impact on the final innovation, or output, node of the phase. Further, as evidence at *t+1* was entered, the innovation node at *t* was also updated, since evidence for a node impacts all ancestor nodes, including innovation nodes; hence, this situation potentially impacted the final output. To remedy this problem, innovation nodes containing no descendants (i.e., having no downstream nodes) were created to better preserve information about the prior time epoch, as shown in Figure 4. These new innovation nodes with no descendants were connected directly to the final output node and were collectively weighted to determine the ultimate output.

Our model assumes that any given product is either innovative or non-innovative, based on the team data that was used to build the model. Another assumption of our model, which was necessary from a model simplification standpoint and will be discussed further, is that activities utilized within time epoch *t* are independent of each other; however, they are dependent on the activities utilized in the prior epoch *t-1*. All three phases were modeled similarly; therefore, only the early phase is illustrated here. The middle and late-phase models each contained one additional node. This leftmost node in the middle-phase model carried information from the early phase (i.e., innovativeness of the early phase). Likewise, the leftmost node in the late-phase model carried information from the middle phase (i.e., innovativeness of the middle phase).

A clustering algorithm was applied to determine empirically-based usage levels for the design categories. Both two-step and K-means clustering algorithms were applied to the activity data, which consisted of the number of occurrences of each design activity by time epoch and team. Since the two-step algorithm yielded more balanced results, it was ultimately selected to perform the clustering. For simplicity, the chosen number of clusters for each design category was three, corresponding to *low, medium*, and *high* usage levels. We desired a small number of categories so that instructors could effectively communicate to their students any changes they should be making to enhance their innovativeness. For example, an instructor might tell a particular team, "You may want to consider doing design communication at a low level at this point and instead invest more time in detailed design by doing it at a high level." Based on the clustering as shown in Table 1, if the number of occurrences for *problem definition* for all team members was less than or equal to three, the team's usage level for *problem definition* was identified as low. If it was between four and eleven, the usage was medium; otherwise its state was high. The clustering for *problem definition* resulted in a cluster solution quality of 0.82, as defined by the silhouette coefficient in the *SPSS* software. *SPSS* identifies cluster solution quality as ranging from poor to good, with "good" extending from 0.50 to 1.00 (Norusis, 2011). Clustering results for the remainder of the design categories are shown in Table 1. Since the category usage levels were determined based on actual student data,

	Usage Level (Occurrences)			
Design Activity	**Low**	**Medium**	**High**	**Cluster Solution Quality**
Problem Definition	0–3	4–11	12 or more	0.82
Conceptual Design	0–1	2–5	6 or more	0.80
Preliminary Design	0–2	3–7	8 or more	0.67
Detailed Design	0–5	6–13	14 or more	0.77
Design Communication	0–2	3–8	9 or more	0.71
Review	0–6	7–17	18 or more	0.70
Management	0–2	3–6	7 or more	0.72
Marketing	0	1	2 or more	0.99

Table 1. Usage Level Definitions for the Design Activities

the definition for what defines low, medium, or high is not necessarily the same across the design categories. This data-driven approach produced results based on the actual nature of the data collected from the students. This approach was necessary, as the usage levels varied greatly by design activity as shown in the table.

Once the design category usage levels were determined, the conditional probabilities could be calculated and entered into GeNie. Since it was assumed that the design categories Y_i in the same time epoch are independent of one another, the conditional probability could be calculated as follows:

$$P\ (Being\ innovative/Y_1 = s_j \cap Y_2 = s_j \cap Y_3 = s_j \cap Y_4 = s_j \cap Y_5 = s_j \cap Y_6 = s_j \cap Y_7 = s_j \cap Y_8 = s_j)$$

$$= P\ (Being\ innovative/Y_1{=}s_j)*P(Being\ innovative/Y_2{=}s_j)*P(Being\ innovative/Y_3{=}s_j)*P(Being\ innovative/Y_4{=}s_j)*P(Being\ innovative/Y_5{=}s_j)*P(Being\ innovative/Y_6{=}s_j)*P(Being\ innovative/Y_7{=}s_j)*P(Being\ innovative/Y_8{=}s_j),$$

where $s_1 \ldots s_3$ represent the utilization levels.

To illustrate the calculation of a conditional probability for a given design category, assume hypothetically that eight *conceptual design* categories were identified as having "low" usage based on the occurrences. Of these, assume five were associated with non-innovative artifacts and three were associated with innovative artifacts. Therefore, this calculation would be performed as follows:

P (Being innovative/conceptual design is "low") = 3/8 = 0.375.

P (Being non-innovative/conceptual design is "low") = 5/8 = 0.625.

This model can ultimately be used to provide feedback to students on the utilization level of each design activity that is most likely associated with an innovative outcome. For example, given that a product is innovative, which usage level for conceptual design (i.e., low, medium, or high) has the highest probability of having occurred and should therefore be emulated? In addition, using the *MAP* (i.e., maximum aposteriori probability) feature within the GeNie decision software, instructors can calculate the most likely retrospective scenario given that a design was innovative and apply this knowledge in the future (Genie Documentation, 2014). For example, given that a design was deemed innovative, the *MAP* algorithm allows a determination of the most likely joint state (i.e., combination) of the utilization levels of the design activities, thereby "telling the story" and allowing a characterization of the design process or scenario for future use.

ASSESSMENT AND VALIDATION RESULTS

In assessing and validating the model, two approaches were used. In the first, the performance of the model using data from the 16 teams that produced either innovative or non-innovative artifacts was determined. Specifically, we compared the probability of being innovative per the model to the actual innovation scores based on the instructors' ratings. The second approach involved determining the model performance using data from the remaining ten teams, which was not used to build the model. It was predicted that these ten teams would have model results in between those of the teams that had produced either innovative or non-innovative products. We also conducted a sensitivity analysis to indicate the robustness of our clustering approach in defining the usage levels of the design activities.

Cross Validation

The model performance results for the 16 teams that produced either innovative or non-innovative products for each of the early, middle, and late phases are shown in Figure 5. For example, in the late phase, the team associated with the highest probability of an innovative product per the model (i.e., 0.970) was also rated as having produced an innovative product by the instructors based upon the assigned project score. In Figure 5, the teams are sorted in descending order based upon the probability of producing an innovative artifact per the model. The horizontal axis corresponds to the various teams, with the left side corresponding to the teams with innovative products and the right side corresponding to the teams with non-innovative products. A threshold model probability value had to be chosen to indicate innovativeness versus non-innovativeness. A probability of 0.50 was chosen as the threshold value; however, different threshold values could have been chosen to do this analysis. Thus, a product was considered "innovative" per the model if the calculated probability of being innovative was greater than 0.50.

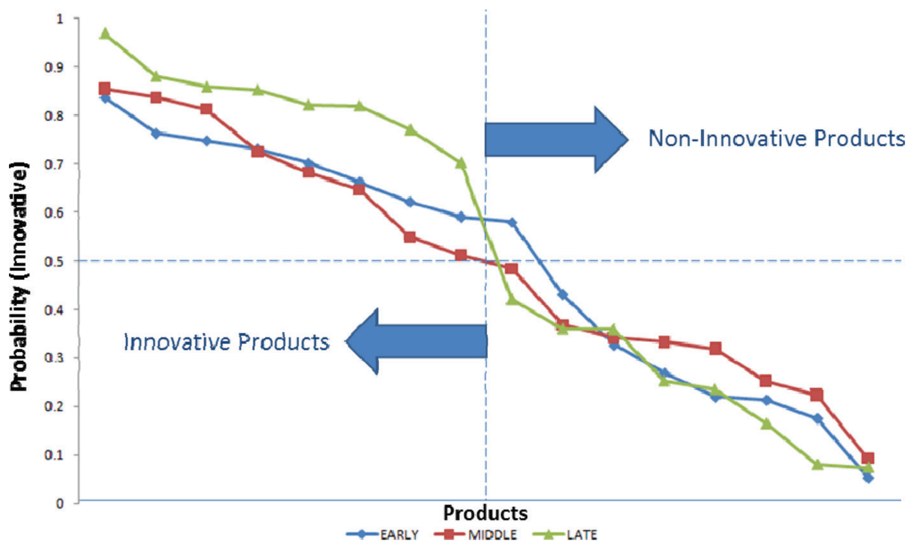

Figure 5. The comparison of the validation results in the early, middle, and late phases.

The accuracy of the model was quite high in the early phase, correctly predicting the innovativeness (or non-innovativeness) of 93.8% of the artifacts (i.e., 15/16). There was a mismatch with only one artifact, but its predicted probability was only 9% above the threshold. The sensitivity of the model was 100% (i.e., all eight innovative products were correctly identified as being innovative by the model), and the specificity was 87.5% (i.e., seven out of the eight non-innovative products were correctly identified as non-innovative by the model). There was also high accuracy for the middle and late phases. The sensitivity and specificity were each 100% for both phases. Thus, the upper left quadrant of Figure 5 represents products that were rated as innovative by the instructors (left side of horizontal axis) and that had a probability of 0.50 or higher (per the model) of being innovative (upper portion of vertical axis). The lower right quadrant represents products that were rated as non-innovative by the instructors and that had a probability of less than 0.50 of being innovative per the model. Figure 5 shows that only one product was not in one of these two quadrants and that our model is therefore accurate.

Further, Figure 5 indicates that the late-phase model produced the strongest results, with innovative products having probabilities closer to one and non-innovative products having probabilities closer to zero, as compared to the early and middle-phase models. Therefore, the late-phase model is considered stochastically dominant relative to the other two models.

The reserved dataset from the 10 teams not included in the development of the model (since their products were neither innovative nor non-innovative) was used as a second validation dataset, and the innovativeness probabilities for the early, middle, and late phases are shown in Figures 6 through 8, respectively. In Figure 6, the innovativeness probability for the 10 teams is approximately in between (albeit with some overlap) that of the teams having non-innovative and innovative products

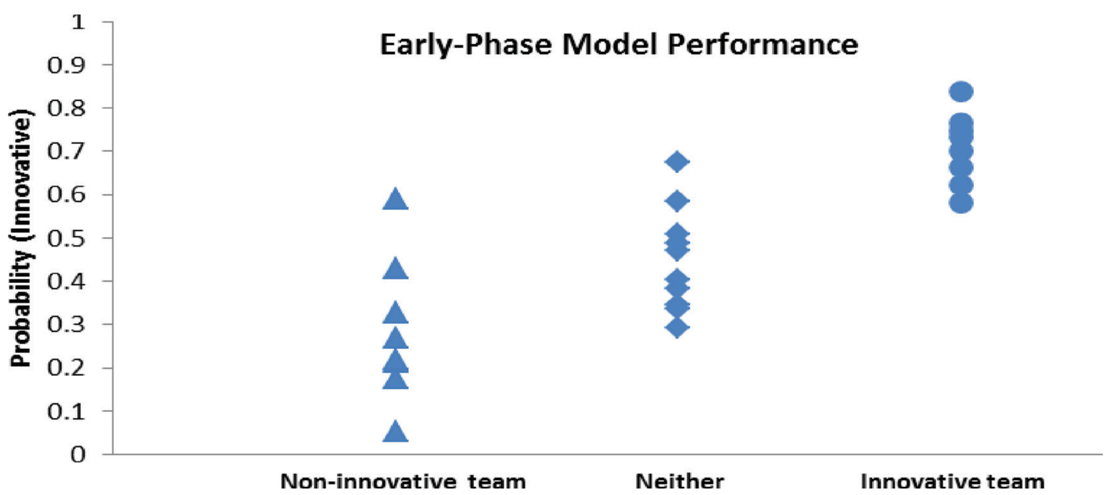

Figure 6. Testing of the Early Phase Bayesian Network Model.

in the early phase. The probability for each of the 10 products is approximately between 0.3 and 0.7. However, among these "middle" rated products, none displayed any strong indication of being innovative or non-innovative. In fact, upon regressing the ascending probability of being innovative on a set of ascending integers (i.e., 1, 2, and 3) representing 1) non-innovative 2) neither innovative nor non-innovative, and 3) innovative products, respectively, the coefficient of determination R^2 was found to be 0.65. From this we concluded that there was a moderate relationship between the innovation rating and the level of innovativeness predicted by the model.

In Figure 7 for the middle phase, all but one of the "neither innovative nor non-innovative" products had an innovativeness probability between 0.4 and 0.6; this product had a probability of 0.19, which was clearly non-innovative. Further, upon regressing the data, the coefficient of determination R^2 remained relatively unchanged compared to the early phase model (R^2 = 0.66).

In Figure 8 for the late phase model, the three groups are more distinct and non-overlapping. The "middle" rated products typically had innovativeness probabilities between 0.4 and 0.6. Interestingly, the probabilities of being innovative were higher in the late phase compared to the middle phase. The coefficient of determination R^2 increased to 0.84. Thus, the late phase Bayesian network model did an exemplary job of correctly predicting the outcome using data from all of the teams, including those whose products were rated in the "middle."

Sensitivity Analysis

Since the utilization levels for the design activities were determined based on clustering of the data, a sensitivity analysis was conducted by changing the upper and lower bounds of the

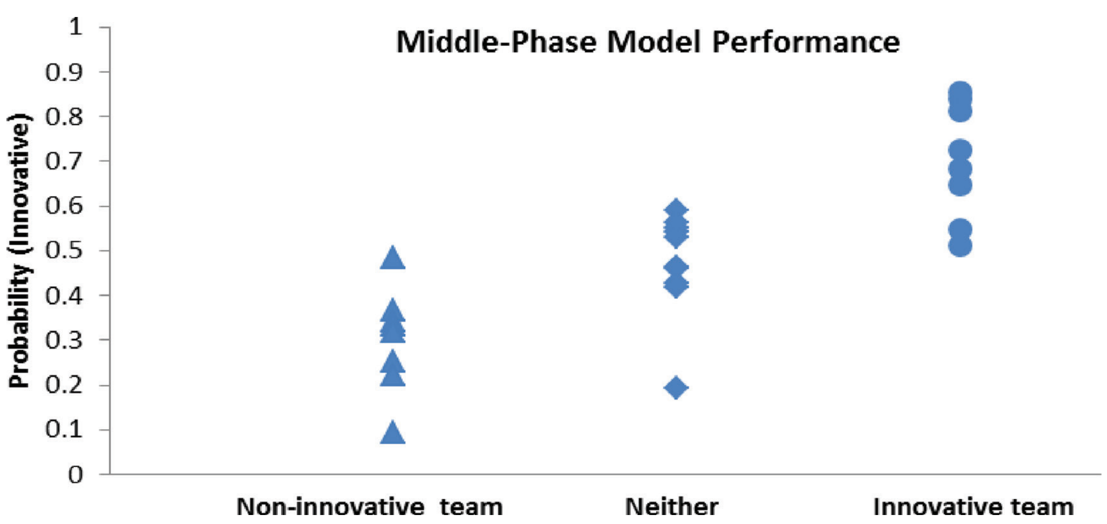

Figure 7. Testing of the Middle Phase Bayesian Network Model.

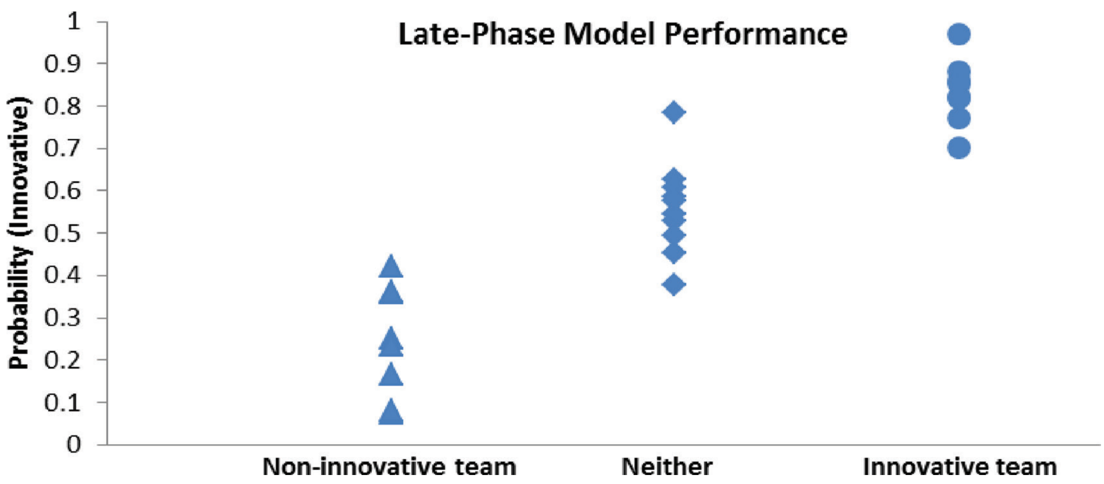

Figure 8. Testing of the Late Phase Bayesian Network Model.

categories. For each variable (i.e., design category), four sensitivity cases were evaluated. Namely, the upper bound of the low utilization category was both increased and decreased by one unit, and the same occurred with the medium utilization category. Note also that when the upper bound of the low category is decreased, the lower bound of the medium category must also decrease by default.

Three innovative and three non-innovative products were selected for the sensitivity analysis. The products were selected based on being potential borderline cases (e.g., an innovative product with a relatively low predicted probability of being innovative such as 0.703). A sensitivity analysis on these six products indicated a fairly robust model, as fluctuations were minimal. Given six products, eight design categories, and four sensitivity cases, there were 192 possible results. Of these, only six results (3%) crossed the threshold from innovative to non-innovative. The *management* category yielded the most fluctuation in that there were three instances in which the probability changed from innovative to non-innovative. With the *conceptual design* category, there were two instances in which the probability changed from innovative to non-innovative. Finally, for the *detailed design* category, there was one instance of a change from innovative to non-innovative.

DISCUSSION AND APPLICATION

Our goal in developing and validating this model was to provide a reliable decision support tool for contributing to the innovativeness of design products. We envisioned that such a tool would be useful to various people involved in product design – engineering educators as well as students. Specifically, we wanted to support instructors in monitoring a team's design process and in providing any needed feedback to the team to assist in a more innovative final product. Thus, we believe this tool can be used as a validated instructional aide for teaching or monitoring innovative design methods, including encouraging students to reflect on the overarching design process and the types of activities helpful to innovation. Both designers and instructors can use the tool to evaluate "what-if" scenarios. Since the GeNie software is freely downloadable, they can experiment with our model and assess the impact of the usage levels of the design activities on the innovativeness of the final product. This tool can make the design process more structured for students by defining clear activities that are based on Dym's design process framework. Our GeNie models are available upon contacting the corresponding author (Renee Clark).

To demonstrate the GeNie development environment and the application of our model, we created two videos as part of this paper. The first video (http://youtu.be/OJuWMprfVDM) provides information on how to use a Bayesian network model within the GeNie decision modeling software. Specifically, it shows how to input known information (or evidence), such as the usage level of a particular design activity, into a model and subsequently obtain output from the model. For our model, the output consists of the probability of being innovative as well as the most probable usage levels of the various design activities. The second video (http://youtu.be/FCbYv3a3ss4) provides information on how engineering educators can use our Bayesian network model to advise their design students. This video was developed using the early phase of our model and demonstrates the switching of a model between innovative and non-innovative settings (using the innovation nodes with no descendants) to uncover contrasts in the most probable usage levels of the design categories, such as *problem definition* or *marketing*. These comparisons and contrasts provide a means to directly advise a design team. For example, if the hypothetical outcome of a product is toggled from non-innovative to innovative using the model and there is an associated change in the most probable usage level of *problem definition* or another activity, valuable insight into recommended process changes can be obtained and used to inform team activities.

CONCLUSIONS AND FUTURE WORK

This paper describes the data-driven development and validation of a Bayesian network model for use as a decision support tool for teaching engineering design. Three separate dynamic models that considered the history of the team's activities were built and validated for each time phase of a multi-term capstone design project. Based on the design activities that a team engages in, and in particular their usage levels, the model allows instructors to predict as well as positively impact the innovativeness of the team's final product. If a team is doing "poorly" in terms of product innovation based on the model, the instructor can advise the team on how to redirect its efforts based upon the activity nodes within the model. The structure of our model was based on Dym's design process framework, with the addition of two product realization categories. Design activity data from capstone teams from two engineering schools were collected over multiple terms and used to parameterize our multi-phase model using a clustering technique. The model was cross-validated using all available outcomes data and was found to be highly accurate (94%-100%, depending on the phase). A sensitivity analysis showed the robustness of our clustering approach in defining the usage levels for the various design activities.

The potential limitations of this work center on three primary areas: size of the data set, generalization of the results to all engineering fields, and the issue of innovative versus "good" design. Regarding the size limitation, the data in this research were from 26 bioengineering capstone teams from two different engineering schools, each covering a different number of terms (two versus three). Of the 26, data from 16 teams was used to develop the model (i.e., eight innovative and eight non-innovative products), with the remainder of the data used for model validation. Although 16 records can be construed as a small data set, the data set is actually rich in information. It reflects 64 students who provided data for approximately 24 weeks, from initial conception to working prototype. Given that each individual student spends approximately 12 hours per week on capstone design work, this data reflects approximately 18,000 hours of design work. As future work, an additional data set would allow for further development and validation of the model across the three time phases.

The ten artifacts that scored in the middle (i.e., neither innovative nor non-innovative) were not used in the development of the model, although they were used in the model's validation. Thus, there is an opportunity to further develop the Bayesian network model to include three outcome states versus two. This research is based solely on data from bioengineering capstone design projects. To make the model more generalizable to other engineering disciplines, future work can incorporate capstone design data from other engineering fields. Also, engineering design experts from multiple fields could be consulted on the face validity of the model using one-on-one interviews. Field testing of the model with engineering design educators is also a necessary future research activity. We surmise that the results of this work are generalizable to design processes that involve a physical artifact. For design processes that result in a process or service, a new study is warranted to determine if similar results are found. Finally, it may also be possible to enhance the model by incorporating the results from the quantitative investigation, with possible adjustments to the models' parameters (Ozaltin et al., 2015).

In summary, the Bayesian network model presented in this paper is a decision support tool for instructors or possibly design managers for monitoring a team's design process and providing feedback as needed. This tool can serve as a data-driven, evidence-based instructional aide for guiding innovative design, including inspiring students to think about the overarching design process and the types of activities that will contribute to their innovativeness.

ACKNOWLEDGEMENTS

This research was supported under an NSF BES RAPD collaborative grant, award number 0602484. In addition, we would like to thank our collaborators Drs. Mark Gartner (University of Pittsburgh), Kay C. Dee, Glen Livesay, and Renee Rogge (Rose-Hulman Institute of Technology) for their invaluable support in our data collection efforts.

REFERENCES

ABET Criteria. (2013). Retrieved from <http://www.abet.org/engineering-criteria-2012–2013/>.

Atman, C., Chimka, J., Bursic, K., & Nachtmann, H. (1999). A comparison of freshman and senior engineering design processes. *Design Studies*, *20(2*), 131–152.

BMEidea. (2014). Retrieved from <http://venturewell.org/students/bmeidea/>.

Callahan, C., Erdmann, J., Hojat, M., Jon Veloski, J., Rattner, S., Nasca, T., & Gonnella, J. (2000). Validity of faculty ratings of students' clinical competence in core clerkships in relation to scores on licensing examinations and supervisors' ratings in residency. *Academic Medicine, 75(10),* S71–S73.

Cross, N. (2001). *Engineering design methods*. West Sussex: John Wiley & Sons Ltd.

Dominick, P., Demel. J., Lawbaugh, W., Freuler, R., Kinzel, G., & Fromm, E. (2001). *Tools and tactics of design*. New York: John Wiley & Sons, Inc.

Dym, C., & Brown, D. (2012). *Engineering design: representation and reasoning (second edition)*. New York: Cambridge University Press.

Dym, C., Agogino, A., Eris, O., Frey, D., & Leifer, L. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, *94*(1), 103–120.

Dym, C., & Little, P. (2004). *Engineering design: a project based introduction (second edition).* New York: John Wiley & Sons, Inc.

French, M. (2010). *Conceptual design for engineers (third edition)*. London: Springer-Verlag London Ltd.

Genie Documentation. (2014). Retrieved from <https://dslpitt.org/genie/wiki/ GeNIe_Documentation>.

Golish, B., Besterfield-Sacre, M., & Shuman, L. (2008). Comparing the innovation processes in academic and corporate settings." *Journal of Product Innovation Management*, *25(1),* 47–62.

Ha, A., & Porteus, E. (1995). Optimal timing of reviews in concurrent design for manufacturability. *Management Science, 41(9),* 1431–1447.

Hyman, B. (2003). *Fundamentals of engineering design (second edition)*. New Jersey: Prentice Hall.

Jensen, F., & Nielsen, T. (2007). *Bayesian networks and decision graphs (second edition)*. New York: Springer.

Krishnan, V., Eppinger, S., & Whitney, D. (1997). A predictive model of sequential iteration in engineering design. *Management Science, 43(4),* 1104–1120.

Lewis, K., Chen, W., & Schmidt, L. (Eds.). (2006). *Decision making in engineering design*. New York: ASME.

Lewis, W., & Samuel, A. (1989). *Fundamentals of engineering design*. New Jersey: Prentice Hall PTR.

McBride, D. (2010). *The process of research in psychology*. Thousand Oaks, CA: SAGE Publications, Inc., 81.

Moskal, B. & Leydens, J. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research and Evaluation, 7(10)*.

Moullec, M., Bouissou, M., Jankovic, M., Bocquet, J., Re´quillard, F., Maas, O., & Forgeot, O. (2013). Toward system architecture generation and performances assessment under uncertainty using bayesian networks. *Journal of Mechanical Design*, *135*.

Norusis, M. (2011). *IBM SPSS Statistics 19 Statistical Procedures Companion*. Upper Saddle River, NJ: Prentice Hall, 396–397.

Okudan, G., & Mohammed, S. (2006). Task gender orientation perceptions by novice designers: implications for engineering design research, teaching and practice. *Design Studies, 27(6*), 723–740.

Ozaltin, N.O. (2012). *The analysis and modeling of the engineering design process: factors leading to innovative outcomes.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Publication No. 3537983)

Ozaltin, N., Besterfield-Sacre, M., Kremer, G., & Shuman, L. (2015). An investigation on the implications of design process phases on artifact novelty. *Journal of Mechanical Design, 137,* 1–12.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K. (2007). *Engineering design: A systematic approach (third edition)*. London: Springer Science + Business Media.

Pugh, S. (1990). *Total design: integrated methods for successful product engineering*. Essex: Addison-Wesley Publishing Company, Inc.

Schumpeter, J. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. New Brunswick, NJ: Transaction Publishers.

Shen, X., Tan, K., & Xie, M. (2000). An integrated approach to innovative product development using Kano's model and QFD. *European Journal of Innovation Management, 3(2),* 91–99.

Simon, H. (1996). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Stempfle, J., & Badke-Schaub, P. (2002). Thinking in design teams – an analysis of team communication. *Design Studies*, *23 (5),* 473–496.

Stiggins, R. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18 (1)*, 23–27.

Yannou, B., Jankovic, M., Leroy, Y., & Okudan Kremer, G. (2013). Observations from radical innovation projects considering the company context. *Journal of Mechanical Design*, *135(2).*

AUTHORS

Nur Özge Özaltin received her Ph.D. from the Industrial Engineering Department at the University of Pittsburgh and her B.S. from Bogazici University in Turkey. Her primary research interest is in improving the innovativeness of students through modeling of the engineering design process. Her methods include Bayesian network modeling and statistical and qualitative data analysis.

Mary Besterfield-Sacre is an Associate Professor and Fulton C. Noss Faculty Fellow in Industrial Engineering at the University of Pittsburgh. She is the Director for the Engineering Education Research Center (EERC) in the Swanson School of Engineering, and serves as a Center Associate for the Learning Research and Development Center. Her principal research is in engineering education assessment, which has been funded by the NSF, Department of Education, Sloan Foundation, Engineering Information Foundation, and NCIIA. Dr. Sacre's current research focuses on three distinct but highly correlated areas – innovative design and entrepreneurship, engineering modeling, and global competency in engineering. She has served as an associate editor for the *JEE*; and is currently associate editor for the *AEE Journal*, and Fellow in the American Society for Engineering Education.

Renee Clark has 23 years of experience as an engineer and analyst. She currently serves as the Director of Assessment for the University of Pittsburgh's Swanson School of Engineering and its Engineering Education Research Center (EERC), where her research focuses on assessment and evaluation of engineering education research projects and initiatives. She has most recently worked for Walgreens as a Senior Data Analyst and General Motors/Delphi Automotive as a Senior Applications Programmer and Manufacturing Quality Engineer. She received her PhD in Industrial Engineering from the University of Pittsburgh and her MS in Mechanical Engineering from Case Western while working for Delphi Automotive. She completed her postdoctoral studies in engineering education at the University of Pittsburgh. Dr. Clark has published articles in the *Journal of Engineering Education*, the *Journal of Engineering Entrepreneurship*, and *Risk Analysis*.