DETC2015- 46537

PATTERNS OF CREATIVE DESIGN: PREDICTING IDEATION FROM PROBLEM FORMULATION

Mahmoud Dinar
Design Automation Laboratory
Mechanical and Aerospace Engineering
Arizona State University
Tempe, AZ, 85287-6106, USA

Yong-Seok Park
Design Automation Laboratory
Mechanical and Aerospace Engineering
Arizona State University
Tempe, AZ, 85287-6106, USA

Jami J. Shah
Professor
Mechanical and Aerospace Engineering
Arizona State University
Tempe, AZ, 85287-6106, USA

Pat Langley
Professor
Computer Science Department
University of Auckland
Auckland, 1142, New Zealand

ABSTRACT

The main objective of our research is to understand the role of problem formulation in creative design ideation. To that end, we have used the web-based testbed of the Problem Map (P-map) computational framework which represents designers' problem formulation in terms of a series of state models, where each state consists of six types of entities in addition to relations within and between different entity types. We gave two design problems to twenty five graduate students in an advanced product design course. We collected their problem formulation data in the P-map testbed and their ideation data through concept sketches. We conducted correlation analysis between variables extracted from the P-maps, and the ideation metrics. We also built regression models for each of the ideation metrics as the dependent variable, and the P-map variables as the independent variables. We used the data from the first problem to predict the ideation scores for the second problem. The predicted results were compared to the actual outcome reported by an independent panel of judges. Models of variety, average and max quality had more accurate predictions while average novelty, average and max quality had statistically more reliable models.

1. INTRODUCTION

Problem formulation is an important part of the design process which has major influences on the design outcome. Yet, this effect is not well understood [1]. The main objective of our research is to understand the role of problem formulation in creative design. To that end, we have developed the Problem Map (P-map) framework [2]. It is a computational framework which facilitates the representation and comparison of designers' problem formulation, in the form of a series of state models. The states show data fragments under six entities (requirements, use scenarios, functions, artifacts, behaviors, and issues), a hierarchical structure within each category (entity type), and relations among the categories. A node-link representation similar to Concept maps [3] provides a convenient way to see a designer's thinking process over time, in addition to comparing different designers approach towards formulating a problem [4]. The formalized and bounded structure of P-map (limited number of categories to choose, and defined relation types) also makes it possible to measure different characteristics of designers' problem formulation, by extracting different types of variables, e.g., the total number of requirements, or the entity type that is mostly expressed in the last quarter of a designing session.

In this paper, we investigate the relation between problem formulation, measured with the aforementioned P-map characteristic variables, and ideation, evaluated by the ideation effectiveness metrics [5]. With data collected from students of a design course working on two design problems, we conduct correlation analyses as well as building linear regression models to reveal patterns of creative designing. Using P-map data as independent variables and ideation metrics as dependent variables in the linear regression model built for the first problem, we also predict the outcomes of the second design problem given its P-map data. We evaluate the accuracy and the precision of our predictions compared to the actual assessments

made by an independent panel of judges. We also locate our results in discoveries from the literature to find similarities or differences. Therefore, we should review the literature first.

2. REVIEW OF PAST WORK

2.1. Relevant findings from empirical studies of conceptual design

A detailed review of relevant findings from empirical studies of conceptual design is beyond the scope of this paper. A recent comprehensive review on empirical studies in designer thinking can be found at [6]. Here, we focus on a few strategies that designers adopt during the early stages of conceptual design: depth or breadth first exploration of the design space, role of abstraction, and order and direction in decomposing the design space. We should state that because of the well-received notion of the co-evolution of the problem and solution spaces [7,8], we take the liberty of not finding a clear distinction between problem and solution spaces; therefore, we interchangeably use the terms design space, and problem-solution spaces.

A few studies that have looked at the differences between expert and novice designers provide examples for the strategies that designers use. Ho [9] have found that expert designers approach directly the main goals and work backward for required knowledge. One can infer a depth-first exploration from this observation. Contrarily, Ball et al. [10] have conducted a protocol study where they have observed experts use more breadth-first search while novices use depth-first search in ideation. However, they also report that experts utilize a strategic knowledge about how to conduct the design process effectively when they face impasses, by switching from a predominantly breadth-first mode of problem solving to an opportunistic depth-first mode. In another protocol study with three subjects Cai, Do, and Zimring [11] have found no relation between creative outcome and depth vs. breadth exploration of the design space.

Another influencing strategic behavior in conceptual design is abstraction. Ward, Patterson, and Sifonis [12] have conducted experiments to investigate the role of abstraction in creative ideation. By actively instructing the participants to formulate the given task in either very specific or more abstract ways, they found that the latter instructions led to more novel ideas. Ball, Ormerod, and Morley [13] have found that experts lean on experiential abstract knowledge while novices rely on case-driven analogies, mainly driven by surface-level cues.

Problem decomposition is another designerly behavior that can affect conceptual design. Liikanen and Perttula [14] have analyzed the prevalence of explicit and implicit problem decomposition modes through a protocol study involving 16 senior students of mechanical engineering. They have found that the subjects implicitly employ top–down problem decomposition while explicit decomposition is rarely used and often does not foster creativity. In contrast, Ho [9] have found

that expert designers are more likely to utilize explicit problem decomposition, leading to more creative ideas.

2.2. Relevant applications of search for patterns, modeling, and prediction

We have already explained how our P-map framework benefits from a node-link representation inspired by concept maps. We have found the work of Oxman on Think maps [15] to be similar to ours in using concept maps. However, it does not exploit the computational framework in search for patterns in the way we do, perhaps because the objective of the framework is strictly comparing students' maps to that of a teacher or a norm. While we have applied P-maps for similar purposes [16], we find more ways in benefiting from such a computational framework, one of which we present in this paper.

One of the established works in using a simpler computational framework compared to concept maps, has been the linkography of Goldschmidt [17] where she has used a node-link representation for capturing the problem space at the fuzzy front end of product design. States in the problem space are described as nodes and operators as links among them. The result is a network that resembles a semantic network, but the structure of a state is not clearly specified; each state represents the problem space at a time with one node. In contrast, our P-map representation shows each state at a time with a network of nodes of different types, adding more specificity and clarity to the representation of the problem space, and in turn facilitating a finer-grained means for comparing different designers' approaches to the formulation of a problem [4].

Cai, Do, and Zimring [11] have developed an extension of linkography in addition to a distance graph to investigate design patterns among designers of different expertise levels and exposure to different stimuli. They modify the definition of links based on lateral transformation and vertical transformation to represent both the breadth and the depth of the problem space explored in design. In lateral transformation the movement is from one idea to an alternative. In a vertical transformation the move is from one idea to a more detailed or elaborated version of the same idea. They report that the more creative the design is, the higher number of alternatives and more chucks and webs are displayed in their representation, the extended linkograph.

A different application of linkography in finding patterns in conceptual design, Kan and Gero [18] conduct protocol studies to acquire information from linkographs. They define two methods to abstract information from the linkographs: one based on clustering, and one based on Shannon's entropy measure. They state that cluster analysis is able to group the linkographs into meaningful clusters, while entropy measures the opportunities for idea development.

There have also been efforts in combining the different representations and search methods that we described above. An example is Hao et al. [19] where they extend previous research on concept map assessment, to develop an evaluation metric in order to predict individuals' problem-solving performance. They

propose their EntropyAvg novelty metric based on the Shannon's entropy in information theory. They have conducted a controlled experiment where they find a strong correlation between individuals' problem-solving performance and their EntropyAvg measure.

3. DATA COLLECTION

3.1. Subjects

Twenty five graduate students in an advanced product design course were the subjects of this study. Since the students had different educational backgrounds, we had them take the Divergent Thinking skill test [20] to have a gauge of conceptual design skill level among our sample subjects. The results suggest a distribution similar to that of 2000 students who had previously taken the test; hence an appropriate representation of individuals' conceptual design skill levels from the sample at hand; see Figure 1. In addition to conceptual design skills, the course syllabus consisted of guidelines for embodiment design (form synthesis, material selection, failure modes), GD&T, model-based design (behavior models such as Modelica, kinematics), and optimization and utility theory. The course prepared the students for taking on two design projects which were assigned in the second half of the semester. The projects are explained in the following section.
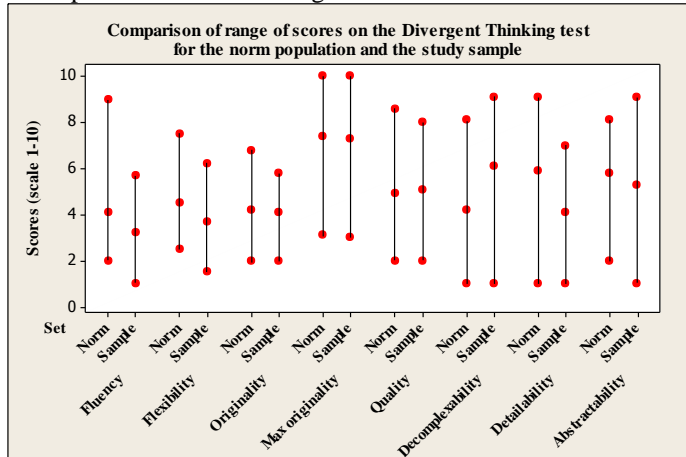

Figure 1. Distribution of the Divergent Thinking test [6] scores among the students in this study compared to a norm population

3.2. Design problems

Two design problems were given to the students. The first design problem—D1—was to design the shot buddy (a design problem from [21]); a device which could return shot basketballs to the shooter, whether the basket was made or missed. The device must be able to automatically adjust the return angle based on the position of the shooter when the ball is shot. The shooter may be initially located anywhere within a 25' semicircle with respect to the basket. It must also accurately and quickly return balls to the shooter and not block the shooters access to the basket. Ideally, the return speed would be

adjustable to accommodate different skill levels. The device should be user friendly for kids ages 10-18, easy to setup and applicable to a wide variety of basket types. The device should be affordable for the average family.

The second design problem—D2—was to design an autonomous surveillance vehicle to automatically and periodically tour the perimeter of two structures, stopping as close as possible to the start point, see Figure 2.
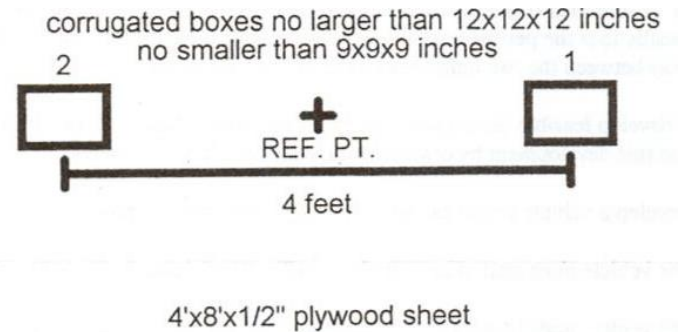

Figure 2. The settings for the D2 challenge problem

3.3. Characteristics of the data set

For each of the two assigned design problems, different types of data were collected corresponding to different steps of the design process. This study involves two steps: problem formulation, and ideation. For problem formulation, students were asked to express their thoughts within the P-map framework, using the Problem Formulator web-based tool [22]. The input data in the tool was short phrases of text put under one of the six entities. These fragments could then be put in a hierarchy within each category, or be related to fragments of other categories. Examples from the data collected for the shot buddy problem (D1) include *requirement* "withstand impacts from the basketball", *use scenario* "outdoor environment" which has the children entities "rain/moisture" and "sunlight", *function* "collect ball" which is related to *behavior* "impact force", *artifact* "funnel", and *issue* "Is there a minimum distance from the net the shooter needs to be able to shoot?".

Different measures can be defined and computed for the collected data. In this paper, we utilize two ways of measuring the problem formulation data. The first one is basically different types of counts of the final state of the P-map; we call these P-map state counts. An example of a P-map state count is the total number of expressed functions. A variety of state count variables can be defined, but in order to avoid the curse of dimensionality, we limit ourselves to the counts of each entity (six variables), in addition to counts of isolated entities (entities in each category that are not part of a hierarchy, i.e., entities with no parents and children; six variables), and counts of disconnected entities (entities that are not related to other categories; six variables).

The second type of P-map variables count traces of specific series of actions which constitute certain problem formulation strategies. An example of a strategy is the tendency to abstract

ideas. When thinking about an aspect of the problem, a designer can either explore details or sub-types of an aspect, or a more general idea about that aspect. The latter ability, abstraction, is considered to influence creative idea generation [12]. Table 1 summarizes five different strategies, whose traces were counted for the collected P-map data.

Table 1. List of formulation strategies traced in P-maps

Strategy	Definition	Conditions
Abstraction	The designer refers to a more abstract aspect of the problem or solution	• Entity parent added at time t1 • Entity child added at time t2 • t1>t2
Entity depth prevalence	The designer develops details of an aspect of the design before looking for its relations to other categories	• Entity parent of type A added at time t1 • Entity child of type A added at time t2 • Entity of type B added at time t3 • Entity of type B related to parent entity of type A at time t4 • t4>t2
Order req_use	The designer follows a specific order adding use scenarios after all related requirements	• A requirement is added at time t1 • A use scenario is related to the requirement at time t2 • Entity of other type added at time t3 and related to the requirement at time t4 • t2<t3<t4
Order req_fun	The designer follows a specific order adding functions after all related requirements	• A requirement is added at time t1 • A function is related to the requirement at time t2 • Entity of other type added at time t3 and related to the requirement at time t4 • t2<t3<t4
Conflict identification	The designer identifies an issue about conflicting requirements	• Requirement R1 is added at time t1 • Requirement R2 is added at time t2 • Issue I1 is added at time t3 • I1 is related to R1 and R2

To evaluate the outcome of the ideation step, sketches were collected from the students and scored with respect to the measures of ideation effectiveness [5]: quantity, variety, average and max novelty, average and max quality. For each design problem, a few high level functions were chosen and judges would identify the solution concepts for each function within each sketch. For the D1 problem, the functions were control-move, collect, store-guide, locate, aim, and return; wheels, funnel, rail, optical sensor, revolving barrel, and spring cannon are examples of each function respectively. For the D2 problem, the functions were move, start, control direction, steer, and stop; propeller, power switch, cam, rudder, and anchor are examples of each function respectively. To reduce bias, different judges evaluated the sketches for the D1 and D2 problems.

4. RESULTS AND ANALYSES

To investigate the relation between problem formulation and ideation, first we conducted correlation analysis between the P-map variables and the ideation metrics. Second, we built regression models with each of the ideation metrics as the dependent variable and the P-map variables as the regressors (independent variables). Third, with the models based on the D1 problem, we predicted the ideation scores for D2. The prediction results were compared to the actual metrics that were computed from the assessment of conceptual design sketches by an independent panel of judges.

4.1. Correlation analysis

A correlation study was conducted to find significant correlations between P-map variables and ideation metrics for problems D1 and D2. For the twenty five data points that were collected from the students (one record per student for each set of formulation input and ideation output) correlation coefficients of magnitude 0.34 (less than -0.34 and more than 0.34) were statistically significant, with 95% confidence. The significant correlations between problem formulation measures and ideation metrics for the D1 and D2 problems are shown in Table 2. Correlations that have a similar trend in both problems are shown Table 3.

The results for problem D1 show a positive correlation between quantity and the number of raised issues, isolated artifacts, and isolated issues. One can infer that leaving artifacts and issues in a flat list, i.e., not focusing on the architecture of the final product or organizing the issues lead to generating more ideas. There is also a high correlation between all five strategies except for entity depth prevalence and quantity; breadth expansion breeds quantity.

Having more issues in a flat list has a moderate positive correlation with variety as well. In addition, the more abstraction and conflict identification happened, the more likely it was for the students to come up with different types of concept solutions. Correlation results for both average and max novelty show that the more the use scenarios were left unorganized, the less the possibility of having original ideas.

Additionally, higher rates of entity depth expansion led to more novel ideas; in other words, the more the students developed an entity before searching for (or being reminded of) related entities in other categories, the more likely it was to propose novel solutions. Finally, students came up with solutions of higher quality when they did consider the relations between issues and other entity types. Best quality of solutions occurred when students did not fail in recognizing the relations between elicited requirements and other entity types, and when they identified conflicting requirements.

Table 2. Significant correlations between problem formulation measures and ideation outcome for D1 and D2

Ideation metric	D1	D2
Quantity	• Isolated artifacts **0.34** • Abstraction **0.40** • Conflict identification **0.43**	• Function **0.38** • Disconnected artifact **-0.40** • Entity depth prevalence **0.53**
Variety	• Abstraction **0.40** • Conflict identification **0.42**	• Order req_use **0.35**
Avg. novelty	• Entity depth prevalence **0.42**	• Disconnected function **-0.41** • Conflict identification **0.42**
Max novelty	• Isolated use scenario **-0.35** • Entity depth prevalence **0.40**	• Disconnected function **-0.38** • Disconnected artifact **-0.37**
Avg. quality	• Disconnected issues **-0.35**	• Behavior **0.38** • Isolated use scenario **-0.35**
Max quality	• Disconnected requirements **-0.48** • Conflict identification **0.40**	

The results for the D2 problem show a few different significant correlations. The total number of identified functions has a moderate positive correlation with variety. The degree to which students made abstractions and found conflicts also have substantial correlations to variety too. An interesting difference between the correlations for the two problems is that entity depth prevalence is positively correlated with average and max novelty. As we will discuss later, the progression of class over time, and the more constrained nature of the second problem resulted in an overall lower variability in the novelty of the students. It is plausible to infer that a more constrained problem requires more focus on each category of entities prior to the designer's shifting attention towards a different category, i.e., within-group depth exploration breeds novelty in more constrained problems. However, this does not contradict with the observation that the more the students failed in organizing the entities within each category and recognizing the relations to

entities in other categories, the worse their ideas were in terms of novelty and quality.

Table 3. Similar formulation-ideation correlations between D1 and D2

Ideation	Measure	D1	D2
Quantity	Issues	0.45	0.36
	Isolated issues	0.45	0.36
	Order req_use	0.67	0.64
	Order req_fun	0.67	0.57
Variety	Issues	0.40	0.43
	Isolated issues	0.40	0.43
Avg. novelty	Isolated use scenarios	-0.35	-0.37

4.2. Regression analysis

To have an understanding of how different variables in the problem formulation together influence ideation metrics, we conducted regression analysis. First, a model was built with P-map state variables as the input variables, and each of the corresponding ideation metrics as the output. Separately, a model was built for the number of times different strategies were utilized during problem formulation with respect to the ideation results. Table 4 and Table 5 show the coefficients of regression for the P-map state variable counts and counts occurrences of strategies respectively.

Significant regressors are shown in bold in Table 4 and Table 5. Since this study is in its early stages, the criterion for choosing significant regressors was set not to be too strict. For P-map state counts, a p value below 0.2 was considered significant; for P-map strategies counts, the bound was set at 0.1. If there were no p values below the set limit, the lowest p value was considered significant (those regressors are starred in the tables). Additionally, regressors that have the same sign in the models for the two problems are italicized. This comparison shows if both problems provide models that can have the same sense with respect to some variables, i.e., if some parts of the models are generalizable and insensitive to the problem. Among the P-map state counts models, average quality has the highest number of variables with similar signs for D1 and D2 (13) while max quality and variety have 4 and 5 similar variables. One can infer that average quality is easier to predict for new problems.

4.3. Prediction results

In order to test how accurate the regression models were in predicting new observations for new problems, data for one problem was taken as a training set and for the other as a test set. We predicted the ideation metrics for the D2 problem based on the regression models of the D1 problem. The structure of the collected data for both problems was the same. For problem formulation, P-map state counts and occurrences of strategies were obtained from the Formulator tool. Concept sketches were assessed by a panel of judges different from the one which assessed sketches of problem D1. For each student, the problem formulation data from D2 was fed into the three models build based on D1. We recorded the difference between the actual

ideation scores, and the predicted outcomes from the three models. Table 6 summarizes how accurately the models for D1 predicted each of the ideation metrics for D2. Since the ideation metrics have normalized scores (on a scale of 1-10), we counted the number of observations which were predicted within 1 or 2 units (10% or 20% margin of error). It can be seen that predictions of variety, average and max quality were highly accurate in models based on state counts and strategies. The strategies counts model is slightly more accurate in predicting quantity, average novelty, and average quality (within 10% margin of error).

Table 4. Comparison of coefficients of regression for the two design problems for the P-map state counts full models; italic: same sign in D1 and D2; bold: statistically significant with P < 0.2; starred: lowest P value above 0.2

Variable	D1 quantity	D2 quantity	D1 variety	D2 variety	D1 average novelty	D2 average novelty	D1 max novelty	D2 max novelty	D1 average quality	D2 average quality	D1 max quality	D2 max quality
Constant	**4.72**	1.85	**5.33**	**3.42**	**5.64**	**2.74**	**7.71**	**4.17**	5.74	3.53	**7.57**	**4.00**
requirement	*0.30**	*0.12*	0.17	-0.27	*-0.02*	*-0.05*	0.09	**-0.18**	*-0.03*	0.15	**0.09**	-0.14
use scenario	-0.26	0.14	*0.07*	***0.86***	**0.31**	-0.11	*0.32*	*0.18*	0.02	-0.29	**-0.18**	0.07
function	*-0.05*	*-0.02*	0.01	0.00	***0.20***	0.03	*0.22*	*0.08*	*-0.05*	*-0.02*	**-0.10**	**0.14***
artifact	0.12	-0.08	*0.01*	0.27	*0.10*	*0.16*	*0.04*	**0.25**	-0.06	-0.05	-0.03	0.05
behavior	-2.26	0.59	-2.55	0.56	*-1.08*	***-0.52****	-2.09	-0.40	0.19	0.27	-0.07	0.32
issue	0.62	**-0.54***	1.35	-0.42	0.60	0.04	**0.92***	-0.13	***0.21****	***0.43****	0.05	0.23
isolated requirement	*-0.07*	*-0.11*	-0.03	0.08	-0.05	0.05	-0.07	0.05	*-0.03*	*-0.01*	-0.04	0.05
isolated use scenario	**-0.60**	0.21	***-1.02***	-0.34	**-0.58**	-0.11	**-0.94**	-0.08	0.07	0.08	-0.10	-0.01
isolated function	*-0.05*	*-0.08*	0.01	-0.17	0.04	-0.01	0.01	-0.06	-0.01	0.03	0.04	-0.11
isolated artifact	*0.07*	*0.27*	0.23	-0.09	-0.09	0.01	-0.04	0.01	*0.11*	*0.07*	*0.11*	*0.03*
isolated behavior	1.66	-0.60	1.60	-0.28	*0.59*	*0.40*	*1.35*	*0.20*	-0.23	-0.40	0.04	-0.37
disconnected requirement	***-0.37***	*-0.18*	-0.24	0.25	-0.01	0.05	-0.14	**0.16**	0.01	-0.11	**-0.11**	0.11
disconnected use scenario	0.20	-0.21	0.11	**-0.95**	-0.13	0.04	*-0.08*	*-0.30*	*0.03*	*0.29*	***0.19***	*0.03*
disconnected function	*0.18*	*0.04*	*0.12*	*0.16*	-0.07	0.02	0.00	0.04	*-0.01*	*-0.09*	0.04	-0.07
disconnected artifact	-0.27	0.35	*-0.39*	*-0.05*	***-0.36***	*-0.09*	*-0.40*	*-0.14*	0.05	0.26	***0.15***	*0.13*
disconnected behavior	0.66	-0.13	**1.12**	-0.57	**0.74**	-0.04	**1.01**	-0.08	*0.14*	*0.10*	0.10	-0.15
disconnected issue	-0.26	**0.95**	-1.12	0.84	**-0.75**	0.11	**-1.05**	0.37	***-0.43***	*-0.48*	**-0.36**	0.12

Table 5. Comparison of coefficients of regression for the two design problems for the P-map strategies counts models; italic: same sign in D1 and D2; bold: statistically significant with P < 0.1; starred: lowest P value above 0.1

Variable	D1 quantity	D2 quantity	D1 variety	D2 variety	D1 average novelty	D2 average novelty	D1 max novelty	D2 max novelty	D1 average quality	D2 average quality	D1 max quality	D2 max quality
Constant	**2.31**	**3.38**	**3.03**	**5.70**	**4.42**	**3.47**	**5.62**	**5.04**	**5.02**	**4.23**	**6.08**	**5.87**
Abstraction	*0.73*	*0.04*	**0.80**	-0.09	*-0.02*	*-0.02*	0.26	-0.06	-0.05	**0.11***	*0.14*	*0.09*
Entity depth prevalence	*0.18*	*0.14*	*0.31*	*0.10*	***0.38***	*0.10*	***0.47***	*0.14*	0.05	0.01	**0.12***	-0.04
Order req_use	**3.67**	-0.20	**4.66**	-1.67	1.06	-0.81	1.77	-0.87	**0.49***	-0.26	*0.62*	*0.40*
Order req_fun	**-1.65**	**1.46***	**-2.42**	**1.70***	-0.67	0.48	-1.05	0.23	-0.25	0.55	*-0.31*	*-0.23*
Conflict identification	*-0.82*	*-2.06*	*-1.02*	*-1.23*	*-0.74*	***-2.29***	*-0.94*	**-2.80**	0.04	0.25	-0.23	**1.15***

Table 6. Accuracy of predicting D2 ideation based on the model built with D1; percentage of accurately predicted scores within 10% and 20% margin of error

Ideation metric	State counts		Strategies counts	
	20% error	10% error	20% error	10% error
Quantity	52	32	60	48
Variety	88	64	76	60
Avg. novelty	76	40	64	48
Max novelty	56	40	60	32
Avg. quality	92	72	96	64
Max quality	92	68	92	76

5. DISCUSSION

This section discusses some of the challenges that we have faced in providing a model that can predict ideation outcome from the characteristics of designers' problem formulation. We also show some statistical measures that we have used to test the reliability of the predicting linear regression models. Additionally, we examine components of the linear models as well as the correlation coefficients that were presented in section 4.1 in order to see if we can find an agreement between our observations and previous findings in the literature.

5.1. Challenges in data collection

We faced three major problems in the way that data was collected. The first challenge related to the difficulties that we experienced in using the data collection tool, the web-based Problem Formulator [22]. Similar to any software tool there is a learning curve. Prior to working on the two problems which we used in this study, the students learned about the tool and its underlying ontology in an hour long workshop, in addition to working on a different practice design problem. Yet, we saw that some students still misused the tool in entering fragments under the wrong categories. Another common mistake was to mistake conjunctive relations with disjunctive relations (which mean alternatives) under a parent node. A part of our ongoing work is to embed a pre-verification system in the tool where users will be prompted to correct their entries, or a more appropriate category is suggested by the tool.

Another challenge in this study was the limitations of selecting appropriate design problems. Even though the ideation metrics have a normalized scoring schema with respect to either a historical pool from previous designs for the same problem, or the sample of designers' concepts at hand, it is difficult to find two problems which lead to ideation outcomes of the same distribution of scores. Some problems, by the inherent constraints that they have, lead to less ideas with less variety in the proposed solutions, which in turn lowers the chance of having high scores of novelty. Figure 3 shows the changes in the variety scores of the students for the D1 and D2 problems. Even though the median remains fairly the same in both problems, the distribution is much narrower in the variety scores of D2 compared to D1. Figure 4 shows how average novelty goes

down from D1 to D2, mainly because the second problem was more constrained since the students were asked to build a working prototype to compete with other students. It is plausible to assume that the students became more conservative in proposing their designs merely due to the fact that they were subconsciously searching for a design that worked.
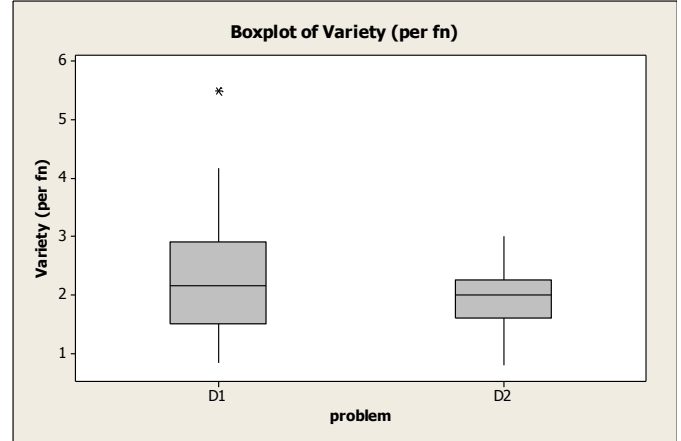

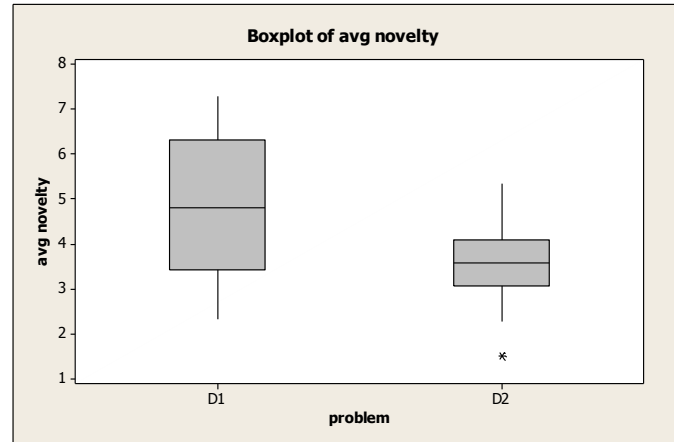**Figure 3. Decreasing variability in variety scores from D1 to D2**

Figure 4. Decreasing mean and variability of average novelty scores from D1 to D2

The challenge in problem selection is coupled with the nature of a progressing class of students throughout a semester in further lowering the variation in outcomes. As the class progressed, through multiple assignments and design projects, the students design skills improved, resulting in a convergence in some of the ideation metrics from D1 to D2. An alternative interpretation of Figure 4 is that it was less likely to come up with a novel idea when students' level of competence had become close.

5.2. Reliability of the results

In order to inspect how reliable the results were, we conducted statistical tests. We have already mentioned in section 4.1 that the correlation coefficients above 0.34 were considered statistically significant for 25 observations with 95% confidence ($p < 0.05$). For the regression models, two statistical

tests were used. One was the test of the hypotheses that the coefficients were 0, i.e., the input variable did not have a linear effect on the output. For this test, the ratio of the coefficient over its standard error, the T statistic, is compared to a Students' t distribution for the same number of observations, and p values are derived. We explained earlier in section 4.2 how we determined which p values to consider.

The other statistic that we checked for each model was the R-squared value. The R-squared of the regression is the degree to which the variation in the output variable is accounted for by the independent variables. The R-squared is often used to examine how accurate the predictability of the model is. Table 7 shows the R-squared values for each of the regression models which were derived for the six corresponding ideation metrics. The test of significance of the model predictability suggests that the prediction results of the P-map state count model is more reliable than the strategies counts model. The results also suggest that average novelty and max quality lead to more reliable predictions in both problems.

Table 7. Test of prediction reliability with R-squared

Predicted variable	State counts		Strategies counts	
	D1	D2	D1	D2
Quantity	65%	75%	64%	33%
Variety	56%	57%	32%	8%
Avg. novelty	78%	65%	30%	25%
Max novelty	66%	72%	24%	35%
Avg. quality	72%	62%	7%	9%
Max quality	87%	71%	19%	7%

The choice of the significant regressors for problem D1 (highlighted in Table 4) was a basis for a preliminary feature selection task. We removed the insignificant regressors and used the P-map state counts reduced model of D1 to predict D2 scores as well. We examined the distribution of the residuals (the difference between the actual ideation scores, and the predicted outcomes) for the prediction which were made with the D1 models. A good model fit, i.e., good predictability of the model would be manifested in a normal distribution with mean zero and a low standard deviation, preferably within 10% of the unit of the metric. Figure 5 through Figure 8 show the distributions of the residuals for four of the ideation metrics which we predicted for D2 in all three regression models based on D1 (full state counts, strategies, and reduced models).

The reduced state count regression model is the best predictor of variety, with the closest mean to zero and the lowest standard deviation. Compared to quantity, predictions of variety (which has the same unit is quantity) in all three models seem to be more accurate. The majority of differences in the actual and predicted variety scores fall close to zero.

Results of predicting average novelty show that the full state counts model has the closest mean to zero; though, the residuals are more spread out compared to the reduced state count model. Another difference between the full model and the other two models is that they undershoot the results, predicting lower variety on average than the actual outcome. The model

based on the occurrences of the strategies has the narrowest spread similarly to the previous metrics.
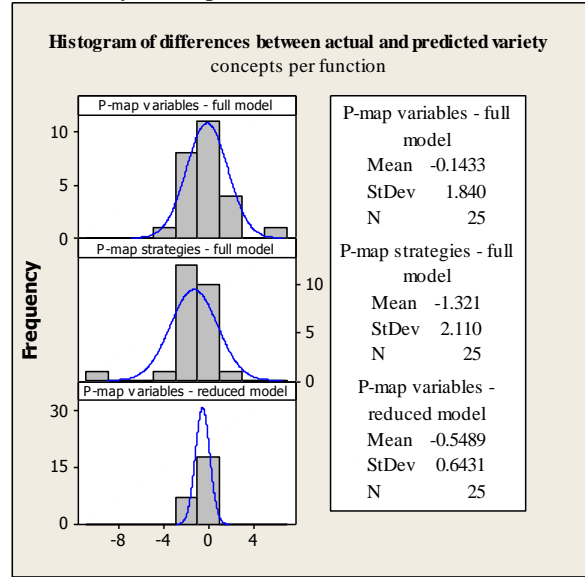

Figure 5. Prediction accuracy of D2 variety from models built based on D1
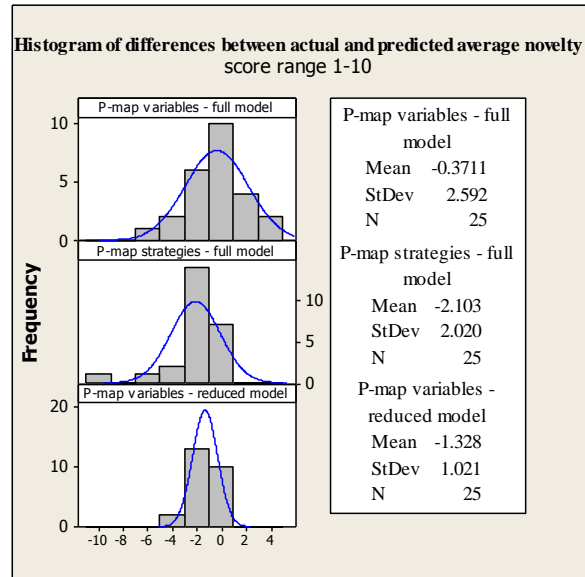

Figure 6. Prediction accuracy of D2 average novelty from models built based on D1

Finally, average and max quality are predicted in an accurate yet imprecise fashion. The distribution of the residuals in Figure 7 and Figure 8 suggest that quality scores are the easiest to get a rough estimate of among all the ideation outcome metrics. All three models result in a mean around zero with nearly identical standard deviations. However, unlike the other metrics, none of the regression models lead to a predictor that one can use with offsetting for an error from the mean.
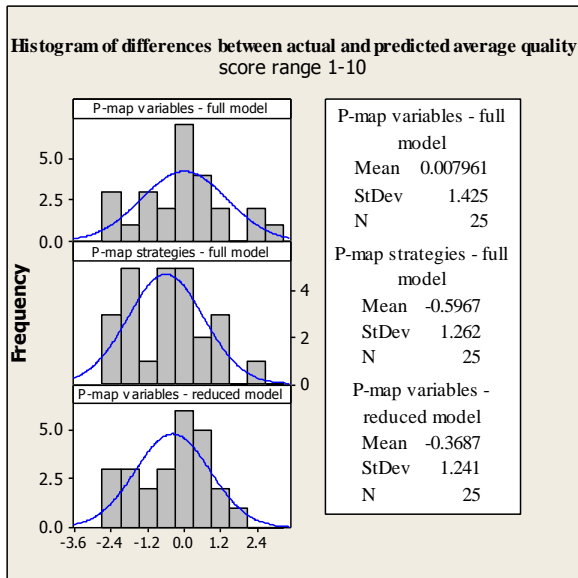
Figure 7. Prediction accuracy of D2 average quality from models built based on D1
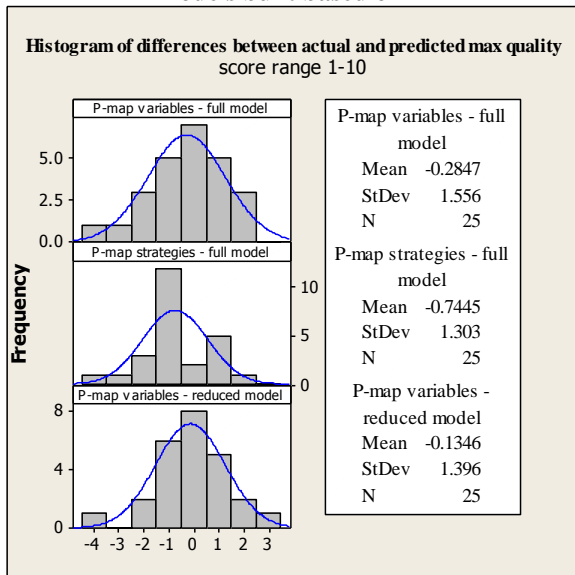

Figure 8. Prediction accuracy of D2 max quality from models built based on D1

We have found abstraction to positively influence the quantity and the variety of ideas, though for one of the two assigned design problems. Our linear regression models also showed a positive but not statistically significant influence on max novelty and max quality (p 0.52 and coefficient of 0.26 for max novelty; p 0.33 and coefficient of 0.14 for max quality). The literature shows that abstraction generally improves novelty [12], and experts utilize abstraction more frequently than novices do [13].

Finally, we have found that following specific orders in decomposing different aspects of the problem breeds the quantity of ideas. We reported quite contrasting observations in the literature where one study has found decomposition does not affect creativity [14] while the other shows that the opposite among expert designers [9].

We should also add that some of the studies that we reviewed in this paper suggest promising alignment between qualitative results with results obtained from quantitative analyses that utilize computational frameworks [11,18]. We believe that using computational methods with data collected on a large scale, coupled with text analysis methods and formal ontologies might help us to reach new findings while avoiding tedious and resource-consuming qualitative research methods.

6. CONCLUSION AND FUTURE WORK

We have used the Problem Map (P-map) computational framework to investigate the relation between problem formulation and ideation in engineering design problems at the early conceptual design stages. Our P-map framework provides the means for extracting different types of information from data fragments that is collected with the web-based Problem Formulator tool. Specifically, two types of variables have been elicited: [final] state counts of six types of entities in addition to the counts of entities that are not related to other entities within each group and between the six groups; counts of traces of strategies which defined with specific series of actions.

We have collected data from twenty five students of a product design course, working on two design problems. Problem formulation data is found in P-maps. Ideation data is in the form of concept sketches which has been turned into the ideation effectiveness metrics: quantity, variety, average and max novelty, average and max quality.

We have conducted correlations and linear regression studies to examine the relationship between problem formulation and ideation. Common correlation results between the two assigned design problems show that the more issues specified in an unorganized way, the higher the quantity and variety of the ideas. By contrast, when use scenarios are not organized in a hierarchical form, there is a slight drop in average novelty. Additionally, developing use scenarios and functions after related requirements significantly influences quantity.

We have built three regression models for each of the ideation metrics: one with all the P-map state count variables, one with the counts of defined strategies, and one with the P-map state count variables that are significant contributors to the initial model (the reduced model). We have built models based on the data from the first problem in order to predict the ideation outcome of the second problem. Our prediction results show that the reduced model is the more accurate in its predictions. However, a test of significance of for model fit suggests that the prediction results of the full P-map state count model is the most reliable among the three regression models. Additionally, we have found average novelty, average quality, and max quality to have more reliable predictor models.

We have faced challenges in evaluating text-based data fragments collected with the Formulator tool, as well as finding appropriate design problems and human subjects that represent

enough variability in the distribution of the ideation outcomes. In the future we plan on collecting more data from more subjects working on different problems in a more controlled experimental setting. Our computation framework facilitates an easily scalable study for a large number of data records, since the data structure is predefined. Once an analysis code or a query is written, it can be used to extract information from the database of the web-based tool at nearly no addition computational cost for a large number of records (different students, problems, and episodes). The way towards higher rate of discoveries might lie in employing computational frameworks for data collection on a large scale, guarded with formal ontologies to assist advanced text and natural language processing, while avoiding resource-consuming research methods such as protocol analysis.

ACKNOWLEDGMENTS

REFERENCES

[1] Harfield, S., 2007, "On design 'problematization': Theorising differences in designed outcomes," Des. Stud., **28**(2), pp. 159–173.

[2] Dinar, M., Shah, J. J., Langley, P., Campana, E., and Hunt, G. R., 2011, "Towards a Formal Representation Model of Problem Formulation in Design," Proceedings of ASME IDETC/CIE, Washington D.C.

[3] Novak, J. D., and Gowin, D. B., 1984, Learning how to learn, Cambridge Univ Pr.

[4] Danielescu, A., Dinar, M., Maclellan, C. J., Shah, J. J., and Langley, P., 2012, "The Structure of Creative Design: What Problem Maps Can Tell Us about Problem Formulation and Creative Designers," Proceedings of ASME IDETC/CIE, Chicago, IL, USA.

[5] Shah, J. J., Smith, S. M., and Vargas-Hernandez, N., 2003, "Metrics for measuring ideation effectiveness," Des. Stud., **24**(2), pp. 111–134.

[6] Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J. S., Smith, S. M., and Hernandez, N. V., 2015, "Empirical Studies of Designer Thinking: Past, Present, and Future," J. Mech. Des., **137**(2), p. 021101.

[7] Maher, M. Lou, Poon, J., and Boulanger, S., 1996, "Formalising Design Exploration as Co-Evolution: A Combined Gene Approach," Advances in Formal Design Methods for CAD: Proceedings of the IFIP WG5.2 Workshop on Formal Design Methods for Computer-Aided Design, June 1995, J.S. Gero, and F. Sudweeks, eds., Springer US, pp. 3–30.

[8] Dorst, K., and Cross, N., 2001, "Creativity in the design process: co-evolution of problem–solution," Des. Stud., **22**(5), pp. 425–437.

[9] Ho, C., 2001, "Some phenomena of problem decomposition strategy for design thinking: differences between novices and experts," Des. Stud., **22**(1), pp. 27–45.

[10] Ball, L. J., St.B.T. Evans, J., Dennis, I., and Ormerod, T. C., 1997, "Problem-solving Strategies and Expertise in Engineering Design," Think. Reason., **3**(4), pp. 247–270.

[11] Cai, H., Do, E. Y.-L., and Zimring, C. M., 2010, "Extended linkography and distance graph in design evaluation: an empirical study of the dual effects of inspiration sources in creative design," Des. Stud., **31**(2), pp. 146–168.

[12] Ward, T. B., Patterson, M. J., and Sifonis, C. M., 2004, "The Role of Specificity and Abstraction in Creative Idea Generation.," Creat. Res. J., **16**(1), pp. 1–9.

[13] Ball, L. J., Ormerod, T. C., and Morley, N. J., 2004, "Spontaneous analogising in engineering design: a comparative analysis of experts and novices," Des. Stud., **25**(5), pp. 495–508.

[14] Liikkanen, L. A., and Perttula, M., 2009, "Exploring problem decomposition in conceptual design among novice designers," Des. Stud., **30**(1), pp. 38–59.

[15] Oxman, R., 2004, "Think-maps: teaching design thinking in design education," Des. Stud., **25**(1), pp. 63–91.

[16] Dinar, M., and Shah, J. J., 2014, "Enhancing Design Problem Formulation Skills for Engineering Design Students," Proceedings of ASME IDETC/CIE, ASME, Buffalo, NY, USA.

[17] Goldschmidt, G., 1997, "Capturing indeterminism: representation in the design problem space," Des. Stud., **18**(4), pp. 441–455.

[18] Kan, J. W. T., and Gero, J. S., 2008, "Acquiring information from linkography in protocol studies of designing," Des. Stud., **29**(4), pp. 315–337.

[19] Hao, J.-X., Chi-Wai Kwok, R., Yiu-Keung Lau, R., and Yan Yu, A., 2010, "Predicting problem-solving performance with concept maps: An information-theoretic approach," Decis. Support Syst., **48**(4), pp. 613–621.

[20] Shah, J. J., Millsap, R. E., Woodward, J., and Smith, S. M., 2012, "Applied Tests of Design Skills—Part 1: Divergent Thinking," J. Mech. Des., **134**(2), p. 021005.

[21] Dieter, G. E., and Schmidt, L. C., 2008, Engineering Design, McGraw-Hill.

[22] Maclellan, C. J., Langley, P., Shah, J. J., and Dinar, M., 2013, "A Computational Aid for Problem Formulation in Early Conceptual Design," J. Comput. Inf. Sci. Eng., **13**(3), p. 031005.