Impact of Mental Iteration on Concept Generation

Pawat Chusilp

Yan Jin[1]
e-mail: yjin@usc.edu

Department of Aerospace and Mechanical Engineering,
University of Southern California,
3650 McClintock Avenue,
OHE-430, Los Angeles, CA 90089-1453

Mental iteration in conceptual design involves repetition of cognitive activities when designers perceive discrepancies of the desired state and current state of design. Although it is believed that mental iteration has significant impact on design process and design results, little proof has been developed and our current understanding of mental iteration is still limited. This paper presents a preliminary study of impact of mental iteration on performance of designers in conceptual design. Mental iteration is modeled as various iterative loops of cognitive activities. An experiment was carried out to study the mental iteration behavior in conceptual design. The analysis of correlation was performed to identify significant associations between design metrics and the number and frequency of different types of iterations. The results provide evidence that mental iteration has not only positive but also negative impact on design performance and different types of iteration yield different impacts. [DOI: 10.1115/1.2118707]

Keywords: mental iteration, cognitive process, conceptual design, protocol analysis, design performance

1 Introduction

Iteration is one of the most basic features of design process. Design concepts emerge and become complete through iteration of analysis, synthesis, and evaluation. Although the term *iteration* is frequently used to describe design process, what people infer by iteration can be very different things. Iteration in engineering design ranges from repetition of design activity to heuristic reasoning processes [1]. Yet, based on what is repeating, one may classify design iteration into two types: iteration of design tasks and iteration of mental activities

For the former, iteration is recognized as repeating of design tasks, which occurs frequently throughout a design project. Iteration of design tasks occurs because the design fails to meet established criteria or new information is obtained since a prior iteration [2,3]. While iteration of tasks is needed for a team of designers to complete a design, it normally leads to delay and variability in production lead time. It was estimated that iteration of design tasks accounted for one third to two thirds of total development time of most product development projects [4]. To date, much research has been carried out [5–7] to prescribe how to manage and reduce unnecessary iteration of design tasks to improve design process.

For the second type of iteration, iteration is recognized as repeating of mental activities of a single designer when he or she is performing a design task. This type of iteration is more difficult to observe because it occurs at cognitive level inside a designer's mind. However, from verbal protocol data, it can be seen that mental iteration occurs frequently throughout design process. Although it is believed that mental iteration can help designers clarify problems, generate ideas, and arrive at better designs, there has been little research on mental iteration, and its behavior, content, and mechanisms have not been clearly understood.

Adams and Atman [8–11] conducted an experiment to study iterative behavior of engineering students. It was found that senior students perform more iterations than freshman students and there is positive correlation between number of iterations and design success. Their work models iteration as transitions between information processing and decision making and identifies specific transition behaviors such as monitor, search, verify, plan, redefine, and capture. The understanding of the roles and patterns of these behaviors helps develop more effective instructional approaches for teaching design and provides helpful references for the development of further research on mental iteration [11].

The focus of this research is mental iteration particularly in conceptual design phase, where the principle of products is determined [12] and approximately three fourths of the final production cost is committed [13]. Our goal is to advance the understanding of mental iteration in conceptual design so that better design methods and supporting tools can be developed. Therefore, instead of approaching mental iteration from a goal-directed problem solving perspective [8–11] and emphasizing behaviors of designers, we approach mental iteration from an idea generation perspective in which conceptual design thinking process is considered as composed of four specific cognitive activities and mental iteration is defined as looping around these activities. We focus on what *ideas* or *contents* are generated, "where" they are generated, how they are enhanced, combined, adopted, or discarded through different iteration loops. This "idea" or "content" focus, in contrast to the "behavior" focus of the work of Adams et al. [8–11], allows us to understand what contents should be captured for *reuse*, and what contents or information should be provided for idea *stimulation* and problem *redefinition*. This understanding is essential for us to develop computer tools for supporting idea generation in conceptual design.

In our previous work [14,15], a cognitive activity model of conceptual design was proposed to capture cognitive activities involved in idea generation and identify different types of mental iteration loops. Protocol studies were conducted to validate the model and study the effect of problems and constraints on mental iteration behavior. The results suggested that each type of iteration loop behaves differently in response to different problem types and constraint conditions. In order to provide recommendations to designers and insights for tool development, we need to understand how different iteration behaviors may lead to different design performance and how skillful designers "manage" their design iterations. If we can understand their impact, we should be able to improve design by encouraging designers to perform good mental iteration behaviors and to devise tools to help them to do so.

In this paper we present a preliminary study of how each iteration loop impacts the designers' performance. This paper is orga-

[1]To whom correspondence should be addressed.

Contributed by Design Theory and Methodology Committee of ASME for publication in the JOURNAL OF MECHANICAL DESIGN. Manuscript received October 12, 2004; final manuscript received April 6, 2005. Assoc. Editor: Clive L. Dym.

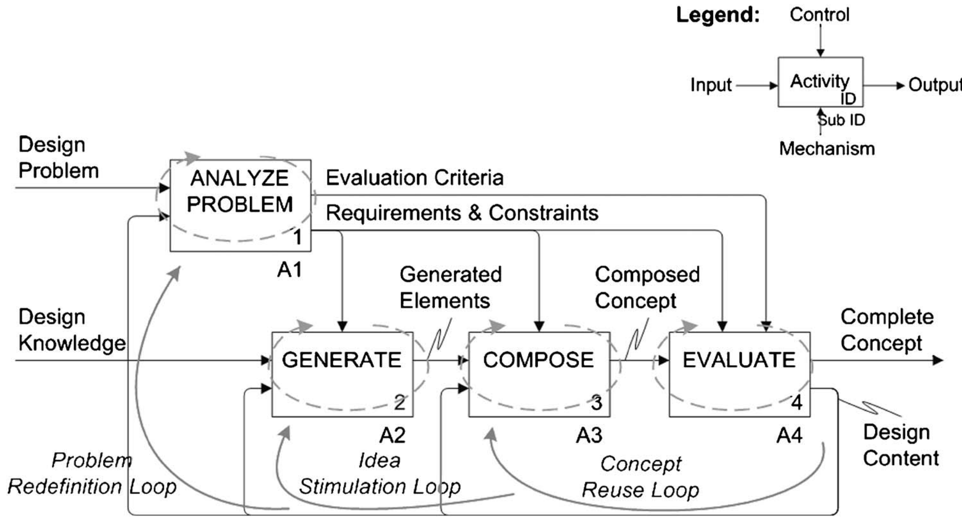

Fig. 1 A cognitive activity model of conceptual design

nized as follows. Section 2 reviews a cognitive activity model of conceptual design from our previous work. Section 3 describes the research method. Results and discussion are presented in Secs. 4 and 5. Finally, Sec. 6 summarizes the paper and draws conclusions. Because our focus is mental iteration, we will use "iteration" and "mental iteration" interchangeably in this paper.

2 Modeling Mental Iteration in Conceptual Design

Conceptual design encompasses various cognitive activities. Mental iteration involves repetition of these activities when designers perceive discrepancies between desired and the current states of design. The iteration can be within one cognitive activity when designers think back and forth before arriving at the desired state of that specific step of design, or it can involve a loop of several cognitive activities when the designers try to achieve the desired design concepts. Mental iteration may be modeled as a sequence of transition between *information processing* activities and *decision* activities [10]. It may also be modeled as various loops, or looping, within and among a number of design-specific cognitive activities [14].

To identify and locate where these iteration loops occur, one needs to understand cognitive activities involved in conceptual design and their structures. A number of cognitive models have been developed to capture the inner working of cognitive processes and structures in conceptual design. These models are mostly based on behavioral observations and cognitive experiments. For example, Ullman et al. [16] developed the task-episode accumulation or TEA model of non-routine mechanical design. Jansson and Smith [17] described a theoretical model of conceptual design process, which describes movement between configuration space and concept space. Schön and Wiggins [18] modeled design process as iterative cycles of seeing-moving-seeing. Finke et al. [19] proposed a model of creative cognition, Geneplore, that can be applied to describe concept generation. Maher [20] and Maher et al. [21] proposed a co-evolution model that describes creative design processes as "co-evolution" between problem space and design space. Shah et al. [22] developed a model of Design Thought Process to describe generation and interpretation of ideas. Cross [23] introduced a general model of creative strategies to describe how highly experienced designers perform creative design tasks. Benami and Jin [24,25] proposed a cognitive model of creative conceptual design that captures interactions between cognitive processes, design entities, and design operations.

Most cognitive models mentioned above treat design process as a single iteration loop. As a result, they provide limited distinction of different types of iteration and their roles and mechanisms in conceptual design. To conduct more insightful study on mental iteration in conceptual design, we need a model that can identify cognitive design activities, addresses information generated from, and used by, these activities, and can capture various types of iteration loops as part of the design process.

2.1 Cognitive Activity Model of Conceptual Design. A cognitive activity model of conceptual design was proposed to study mental iteration in conceptual design [14]. The model was constructed based on a cognitive model of creative conceptual design [24,25] but we take a *process view*, rather than an *interaction view* [25] to model mental iteration. Furthermore, IDEF0, a schematic language designed to describe functions and processes [26], was applied to represent cognitive activities in terms of their relations. In IDEF0, every activity has four types of interfaces, *inputs* (on the left side), *outputs* (on the right side), *controls* (on the top), and *mechanisms* (on the bottom). The differentiation of the four interfaces allows us to explore relations between mental activities and identify the roles of various contents in mental iteration with respect to different phases of thinking process in conceptual design.

Figure 1 presents the overview of the model, which comprises four key cognitive activities of idea generation process, i.e., *analyze problem*, *generate*, *compose*, and *evaluate*. These key cognitive activities are conceived from *generative processes* and *exploratory processes* described in Benami [25] and Finke et al. [19]. They can be further decomposed into subactivities at a more detailed level. But the details and their relations are difficult to clarify so we do not include the detail structure here. However, one of the possible patterns can be found in Chusilp and Jin [14]. In the following we briefly describe the major activities and their subactivities.

2.1.1 Analyze Problem. *Analyze problem* activity is carried out by designers to develop understanding of the problem on hand and explore requirements and constraints that must be satisfied and maintained by the design. Through problem analysis, design goals are set and constraints, and requirements are defined. During design, the problem definition may be elaborated or revised later, and the definition change will result in changes in constraints and requirements. As part of problem analysis, solution criteria are also determined from design goals, as indicated in Cross's general model of creative strategies [23]. *Analyze problem* may comprise the following subactivities: *formulate design goals, establish criteria, determine requirements and constraints*, and *redefine problem*.

2.1.2 Generate. The *generate* activity is where designers generate design ideas. Given problem requirements and constraints, designers retrieve from their memories relevant information and knowledge to create initial design ideas. Based on the Geneplore model [19] and our previous work on cognitive modeling of creative design [24,25], we include both memory retrieval and perceptual stimulation as part of the *generate* activity. “Perceive” acts in response to iteration that stimulates the designer’s ideation. Therefore, the subactivities of generate may include *generate* are *perceive* and *retrieve*.

2.1.3 Compose. The *compose* activity is introduced to capture the evolution of initial design ideas into identifiable design concepts [25,27,28]. This activity is performed when designers combine new ideas generated from their mind with the ideas and/or concepts generated from previous iteration cycles. The combined ideas are then further transformed into more matured design concepts. Although many models treated “compose” as part of “generate” [19,25], differentiating the two provides opportunities for us to study how iteration interacts with idea generation and evolution. This key activity may comprise the following subactivities: *associate* and *transform*.

2.1.4 Evaluate. Once a concept is formed, it is then *evaluated* against design requirements, constraints and criteria. As an exploratory cognitive process [19], “evaluate” is performed by designers to make sure a generated design concept is relevant, useful, and good. Relevance and usefulness of a concept are determined against design requirements and constraints, while goodness depends on design criteria. The *evaluate* activity could comprise the following subactivities: *examine* and *select*.

As shown in Fig. 1, designers analyze the problem to identify and define design requirements, constraints, and evaluation criteria, which are used to guide other activities in design. Initial design ideas are generated as pre-inventive structures by designers based on their background knowledge and recognition of the current design context. Identifiable ideas are then composed to form design concepts. These composed ideas are evaluated to determine if they are complete and satisfactory. Designers may discover new conflicts inside evaluated ideas that make them go back to either redefine the original design problem or identify subproblems. The perception of the intermediate ideas and concepts may also stimulate the designer to generate other ideas. Moreover, the generated ideas may be reused as they are combined with other ideas and when design situations change. The process also involves local iterations within each major activity.

2.2 Iteration Loops of Mental Activities. The proposed cognitive activity model of conceptual design identifies various types of iteration loops embedded in the idea generation process. The loops among key cognitive activities are, namely, *problem redefinition* loop, *idea stimulation* loop, and *concept reuse* loop. We call these loops *global iteration* loops.

In the *problem redefinition* loop (PR loop), designers realize the need for revising the original problem definition after evaluating the composed design concepts against the problem requirements and constraints. They can choose to modify the current problem definition or decompose the problem into subproblems. The problem redefinition often leads to the change of requirements and constraints. It can be expected that the problem redefinition loop facilitates idea generation by allowing expansion of problem space and co-evolution of problem space and solution space [21].

Through the *idea stimulation* loop (IS loop), generated concepts stimulate designers to generate other ideas. Our previous research has shown that there exist patterns of stimulation in which certain types of intermediate design concepts, e.g., physical behavior concept, appear to be more effective in stimulating idea generation [24,25]. It can be expected that the idea stimulation looping may increase the number of newly created ideas.

Through the *concept reuse* loop (CR loop), designers pick up previously generated ideas and then reuse them to compose new design ideas. It can be expected that the concept reuse looping can increase opportunities to better use created ideas.

Besides global iterations, there are *local iteration* loops within each major cognitive activity. Local iterations are carried out by designers to evolve or evaluate ideas until the “desired state” of the ideas is reached so that the design can be moved to the next step.

While the model in Fig. 1 is a cognitive activity model of conceptual design process, not design iteration per se, it does show what kinds of iterations are possible, where iterations occur in the design process and how the iterations might interact with the cognitive activities and impact the overall design process and performance.

2.3 Comparison to Other Models. The cognitive activity model of conceptual design described above was conceived from our previous research [24,25] and influenced by other research as some features of the model resemble those of others’ models (e.g., iteration of problem and solution, which is described as co-evolution between problem space and solution space [29], iteration of problem definition [30], feedback from evaluation stage back to generation stage [31–33]). However, our model describes the conceptual design process in terms of four specific *cognitive activities* and mental *iteration loops*. This integrated representation opened possibilities for us to explore how design mental iterations can impact on design process and design performance.

In addition, it is interesting that though we take a different approach from Adams and Atman [8–11] to identify and classify iteration loops, there are similarities. In their work, types of iterative design “cycles” are obtained from critical literature review of research in design and defined by activities that are involved in the transition. Their iterative design cycles include problem scoping, solution revision, coupled problem scoping and solution revision, and self-monitored cycles. Nevertheless, there are differences. Our model includes idea stimulation loop and concept reuse loop that capture the idea generation phenomenon in conceptual design. On the other hand, the self-monitored cycle does not explicitly appear in our model but compounded with all types of iteration loops, more or less, as a part of the control of the cognitive activities.

3 Research Method

To study the impact of mental iteration on design performance, we first need to conduct protocol studies to measure designers’ mental iteration behavior (i.e., numbers of each loop, percentages of each loop, etc.) and their performance. Next, correlation analysis needs to be carried out to test the associations between the measures of mental iteration and design performance.

3.1 Protocol Studies. The experiment comprised protocol studies of 16 subjects working on a design problem. Two primary methods of protocol analysis are retrospective reports, where subjects are asked questions of what they thought after they finish the task, and concurrent verbal reports or think-aloud method, where subjects have to speak what they think out loud during the experiment [34]. The think-aloud method was selected because it can reveal the sequence of subjects’ thinking process. However, the think-aloud method also has the disadvantage that subjects are often neither familiar nor comfortable with speaking out loud while thinking. Therefore, it is necessary to train the subjects to become more familiar with the method in order to minimize the influence of this disadvantage.

All subjects in our experiment were 2003–2004 academic year students at the University of Southern California. Fourteen of them majored in mechanical engineering and two in industrial and systems engineering. All subjects had limited work experience and little experience of skiing. Thirteen of them were male while the other three were female. Participation was on a voluntary basis. Subjects were tested individually in a quiet room to prevent distraction. Before starting the experimental problem, a brief think-aloud instruction with an approximate 30 min warm-up task was given to the subjects. There was no time limitation but most subjects finished the problem within 1 h. The whole experiment sessions were video taped by two cameras: one from the top to capture their sketches and the other from the front to capture their gestures.

The selected problem for the experiment was required to meet a few requirements. First, the subjects were students so it should not require more than basic engineering education to develop solutions to the problem. Second, the design space should be open enough for a wide variety of solutions. Third, the problem should not be too common so that subjects generate original ideas. For this experiment, the problem was stated as follows:

"*Today ski and snowboard are widely used as personal transportation tools on snow. But, to be able to use them, a lot of skill and experience is required that normally a user cannot learn within one day. Moreover, ski and snowboard cannot run uphill because they are moved by the gravity. Your task is to design other options of personal tools for transporting on snow. The design must be human-powered (powered by the user himself or herself) so that it can run without help from engine or gravity. The design must allow user to control direction and brake. In addition, it shouldn't require much time to learn how to use it.*"

The analysis of protocol followed the following steps. First, verbal protocol recorded from entire design sessions was transcribed. The next step was activity matching. The transcripts were divided into segments, in which each segment expresses the subject performing a cognitive activity. Then the four interfaces of each activity were identified and encoded using a formal language [15]. Finally, the numbers of iterations in each loop are counted. The details of code definition and illustration of encoding process were presented in Jin and Chusilp [15]. Two operators were involved in the coding process and the average numbers of loops were used. The consistency between two operators was about 90% in activity matching, 75% in identifying interfaces, and 95% in counting loops. Overall, we obtained 16 sampling points.

3.2 Measurement of Design Performance.

3.2.1 Existing Approaches to Design Evaluation. Before carrying out this empirical study, we need to define how we measure design performance. One may assess design performance by evaluating outcomes of the design process, which are design concepts or solutions. In design research, there is no common method to evaluate design concepts. Different researchers often applied different methods to evaluate design concepts in their studies. For example, in the evaluation of wall-mounting design in Fricke [35], an evaluation method, which is referred to as VDI-guideline 2225 [36], was used. Solutions were evaluated to a set of requirements derived by experts, which comprehensively covered the problem qualification profiles. The score for each requirement was rated from 0—requirement is met inadequately to 4—requirement is met optimally. Then the final score was calculated from the average of all scores of each requirement. The assessment was done independently by two groups of experts.

In the litter disposal design experiment, which appeared in Dorst [29,37], quality of the design concepts was determined by the following measures: *ergonomics, technical aspects, aesthetics, business aspects, creativity*, and *total judgment*. The *total judgment* was not another mean of the other scores, but a separate overall impression score. The scores, ranging from 0 (poor) to 10 (excellent), were given based on the judgment of five assessors. Alpha coefficients for the agreement between the judges were checked for inter-rater reliability.

In Atman et al. [38], two measures were used to evaluate a playground design. The first was a measure of whether the design met the constraints specified in problem statement. The second was a quality score, ranging from 0 to 1, based on fulfillment of specific criteria, diversity, aesthetics, injury protection, uniqueness, and technical feasibility. The evaluation was done by two assessors separately to maintain the reliability.

Dahl et al. [39] evaluated a car jack design for seniors by *originality, usefulness, and appeal constructs*. Scores for three measures were determined by three-item, seven-point semantic differential scale [40–42]. The three items composing the *originality* scale were unique (7)—ordinary (1), original (7)—commonplace (1), and fresh (7)—routine (1). The three items for *usefulness* are useful (7)—useless (1), effective (7)—ineffective (1), and worthwhile (7)—worthless (1). The *customer appeal* was composed of appealing (7)—unappealing (1), likeable (7)—not likable (1), and desirable (7)—undesirable (1). The evaluation for each measure was done separately by a sample group of seniors, the customer target of the product. There were 14 seniors evaluating *originality*, 16 for *useful*, and 12 for *appeal*.

A more formal approach was proposed by Shah et al. [43]. Four metrics were proposed for evaluating outcomes of idea generation methods. They are *novelty, quality, variety*, and *quantity*. The score of each metric, except *quantity*, is in the scale of 0 to 10 with 0=poor and 10=good. Much of calculation of scores are based on the formula that makes their method less dependent on assessors' judgment and reduces bias.

In general, there are three issues to concern in the evaluation: the measures (i.e., usefulness, ergonomics, quantity, etc.), the procedures to evaluate or determine scores of measures, and the assessors who carry out the evaluation. The measures should be chosen based on the product being designed in the experiment because a measure may be suitable for one design but not another. The evaluation procedure should be as systematic as possible to reduce individual bias and maintain consistency. The assessors should have good knowledge of the product from the perspectives of the measures. Assessors can be a group of selected customers if the measures are related to customer satisfaction or product appeal. On the other hand, assessors have to be product designers or experts if the measures are technical. In addition, the evaluation should be carried out by more than one assessor to obtain reliable results and avoid individual bias. A method with a systematic procedure may require fewer assessors than a method that heavily relies on peer assessment.

3.2.2 Design Metrics. We found that the design metrics in Shah et al. [43] measure important creativity aspects in which we are interested. The metrics are not problem specific so they can be used in other design problems in our future experiments. Furthermore, the method provides more objective procedures to calculate each metric and helps in reducing inconsistency and bias. Finally, this method has been applied by other researchers [44] to measure ideas generated in the experiment. Thus we decided to use their method for evaluating generated ideas in our experiment. A small variation from the original work was introduced to fit the needs of our study. For single stage evaluation, the formulas to calculate four design metrics are described as follows and an example of how scores are calculated is presented in Sec. 3.2.3.

3.2.2.1 Novelty. *Novelty* is a measure of how unusual or unexpected an idea is. It is remarked that not every novel idea is a good idea. The novelty score can be computed from Eq. (1).

$$\textit{Overall novelty score} = \sum_{j=1}^{n} f_j S_j \qquad (1)$$

where

f_j=*The weight assigned according to the importance of Function j*

S_j=*The novelty subscore for Function j*, *ranging from 0 to 10*

Two approaches were suggested to calculate S_j: (1) from assessors' *priori* knowledge or (2) from Eq. (2).

$$S_j = 10 \times (T_j - C_j)/T_j \qquad (2)$$

where

T_j=*The total number of ideas for Function j*

Table 1 Novelty subscore rating

Subscore	$f1$ Propulsion	$f2$ Control Direction	$f3$ Brake
3	Paddles with wheel(s), Sketching	Handles/Steer with front wheel(s)	Cramp brake
7	Paddles with threads	Handles/Steer with ski pad(s)	Lever brake, Snow plug
10	Others	Others	Others

C_j=*The number of the current solutions for Function j*

In our experiment, we use the first approach. For a self-powered personal transporter on snow design, there are three main functions (m=3): propulsion (f_1), brake (f_2), and control direction (f_3). Because the first function is the most important, the weights are assigned as follows: f_1=0.50, f_2=0.25, f_3=0.25. The novelty subscore S_{1jk} for each function is rated based on features presented in Table 1.

3.2.2.2 Variety. *Variety* is a measure of how well one has explored solution space. High variety of generated ideas indicated more probability of finding better ideas in the other area of solution. The variety score can be calculated from Eq. (3).

$$\textit{Overall variety score} = \sum_{j=1}^{n} f_j S_j \tag{3}$$

where

f_j=*The weight assigned according to the importance of Function j*

S_j=*The variety subscore for Function j*

The variety subscore can be computed from Eq. (4).

$$S_j = 10 \times \sum_{k=1}^{m} s_k b_k / M_{\max} \tag{4}$$

where

s_k=*The predefined weight of level k in the genealogy tree*

b_k=*The number of branches at level k in the genealogy tree*

$M_{\max}$=*The maximum possible overall variety score*

Genealogy trees of idea generation must be constructed and inspected to evaluate variety score. The suggested values of s_k are 10, 6, 3, 1 for physical principles level (s_1), working principles level (s_2), embodiment level (s_3), and detail level (s_4). For our study, we consider three levels of genealogy tree, which are working principles level, embodiment level, and detail level. So $M_{\max}$ becomes the multiplication of the number of design concepts and the weight of the highest abstraction level of a genealogy tree, which is the working principles level for our experiment. Same set of functions and weights for novelty are used.

3.2.2.3 Quantity. *Quantity*, as described in Shah et al. [43], is the total number of ideas generated by a subject over a design session. Though most subjects in our experiment produced just one final design concept, they developed several others but discarded them during the design process. Thus, instead of counting just one final concept, we consider the total number of both selected and unselected ideas for each function that were generated during the design session. This count can be, more or less, determined by counting a number of leaf nodes in a genealogy tree of ideation generation for each function. The overall quantity score is obtained from the weighed average of subquantity scores of each function as shown in Eq. (5).

$$\textit{Overall quantity score} = \sum_{j=1}^{n} f_j S_j \tag{5}$$

where

f_j=*The weight of Function j*

S_j=*The quantity subscore of Function j* The subquantity score of Function j can be calculated from Eq. (6). The multiplication of 10 is for normalizing purpose.

$$S_j = 10 \times \frac{N_j}{N_{j\max}} \tag{6}$$

where

N_j=*Count of leaf nodes in genealogy tree for Function* f_j

$N_{j\max}$=*Maximum* N_j *in all subjects for Function* f_j

From the formula, the one who generates the maximum number of ideas for every function receives the highest quantity score. The same set of functions and weights for novelty is used. Because the quantity score expresses only the number of ideas that are created but how much each idea is different, it is recommended to view quantity score together with variety score so that we can have a better view of how many ideas each subject produced and how much they are different from each other.

3.2.2.4 Quality. *Quality* is a measure of how good an idea is based on technical aspect and how close it comes to meet requirements and constraints. The quality score can be computed from Eq. (7)

$$\textit{Overall quality score} = \sum_{j=1}^{n} f_j S_j \tag{7}$$

where

f_j=*weight of Function j*

S_j=*Quality subscore of Function j*

Same set of functions and weights for novelty is used. S_{jk} for each function are rated according to following rates: 10 if feasible and good performance, 7 if feasible, 3 if infeasible.

3.2.2.5 Best novelty versus final novelty. In our experiments, we observed that many subjects discovered novel ideas but they later dismissed them. As a result, their final design ideas at the end of design session may not be novel although they had generated more novel ones. To study how novel ideas are generated and dismissed through iteration, we evaluate the novelty of both final ideas and those generated during design session. For convenience, we call the novelty of a selected final idea *final novelty* and that of the best idea a subject had explored *best novelty*. Note that because the score is composed of subscores of three primary functions, *best novelty* score may be composed of novelty subscores of each function from different ideas.

For quantity and variety, the score is evaluated from number of ideas and number of branches in a genealogy tree. So there is no *best* and *final*. For quality, a final design concept is expected to be more complete and more feasible so the *best* quality and *final* quality should always be equal. It is possible that ideas generated during design session could be better than a final design. However, in our experiment, *best* quality and *final* quality are all equal.

3.2.3 Illustration. In the following, we demonstrate the evaluation of design scores for a subject in our experiment. The design session is briefly described.

During the design, this subject first analyzed requirements of the problem and itemized functional requirements. Then he explored many choices of propulsion system including an unusual

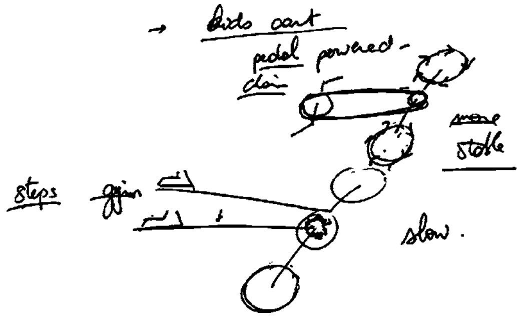

Fig. 2 Design sketch of working principles

idea, i.e., a stepping mechanism, but this idea was discarded as the designer realized that it was not effective. Using analogy of bicycle propulsion, one of the ideas the designer created was a paddling mechanism. The subject then adopted a bicycle for developing his design. The subject intended to have his design to be stable and does not flip over easily so he included two rear wheels and designed the wheels to be thick enough to prevent sinking into snow. As the subject realized the slippery ground surface problem, needles were implanted into the wheel surface to increase the friction between the wheels and snow. After finishing the propulsion system, he started to think about how to control the direction and how to brake. The use of paddling mechanism from bicycle seemed to fix the subject to explore his choices of controlling direction only in variations of bicycle steering column and handle bar. For the brake system, the subject explored several choices and finally chose a braking stick design that directly applies pressure toward the rear wheels by a stick. Then he realized the need to propagate the movement from hand to the braking stick so he designed braking lever and a pivot that user can pull and the sling are wired through the frame of the system and connected to the braking stick. The final design concept created by

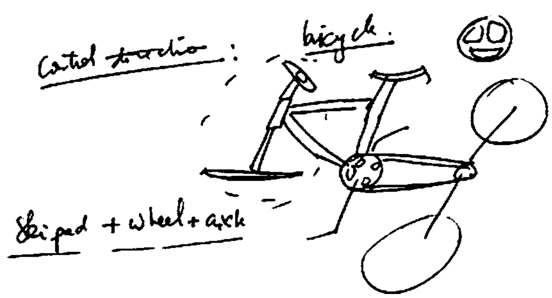

Fig. 3 Design sketch of the finished idea

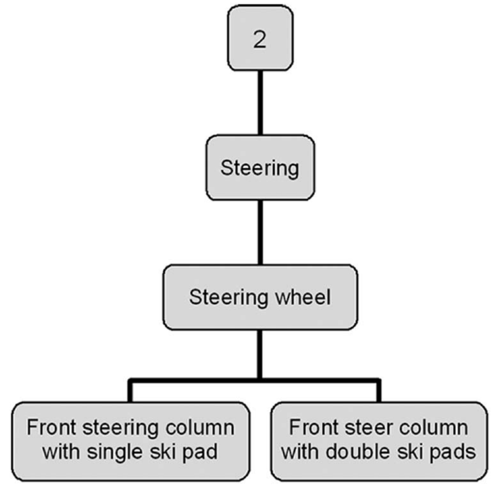

Fig. 5 Genealogy tree for function 2—control direction

this subject was a tricycle-like vehicle that has a steering wheel with front ski pad connected to the front steering column and two rear wheels with a lever-stick braking mechanism.

Figure 2 presents a portion of the original sketch drawn by a subject that illustrates the ideas created to satisfy Function 1, the propulsion. Figure 3 shows a portion of the original sketch that illustrates the finished concept.

Genealogy trees of idea generation were constructed based on subjects' protocol data and original sketches. The number inside the box at the top of the tree denotes the total count of generated ideas that is determined from the number of leaf nodes in the tree. Figures 4–6 show subject's genealogy trees for function 1—propulsion, function 2—control direction, and function 3—brake.

These genealogy trees are used to calculate the quantity and variety scores. As we discussed earlier, the primary function attributes of a self-powered personal transporter on snow are (1) propulsion, (2) control direction, and (3) brake. The weights for each function are assigned as follows:

f_1=*Weight of Function 1—Propulsion*=0.50
f_2=*Weight of Function 2—Control Direction*=0.25
f_3=*Weight of Function 3—Brake*=0.25

3.2.3.1 Novelty evaluation. Let S_1, S_2, and S_3 be novelty scores for Functions 1, 2, and 3. The values of S_1, S_2, and S_3 are given based on the features presented in Table 1. For *final* novelty, $S_1=6$ as a paddling mechanism with two rear wheels was used, $S_2=7$ as steering wheel with front guiding ski pad was used, S

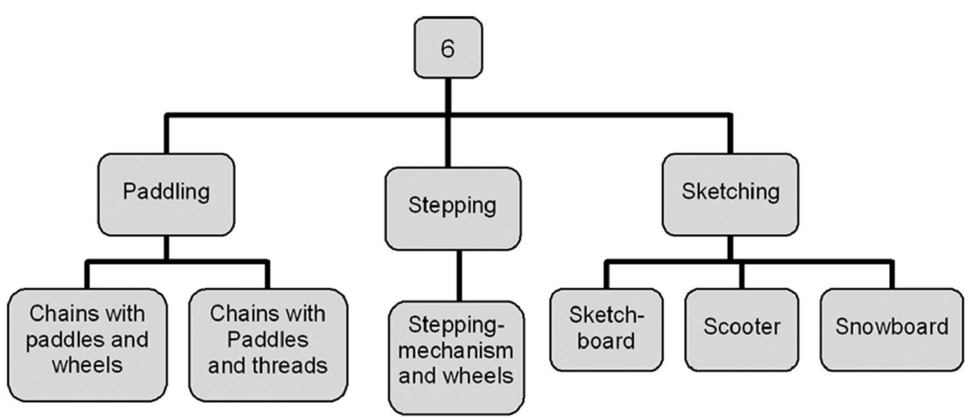

Fig. 4 Genealogy tree for function 1—propulsion

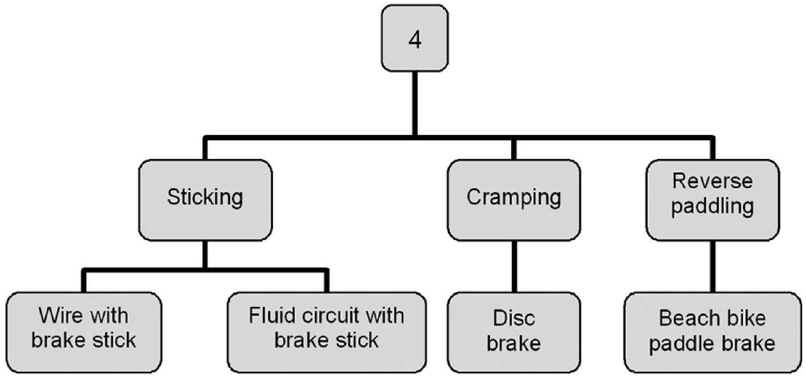

Fig. 6 Genealogy tree for function 3—brake

=3 as the wheel brake was used. The *final* novelty score is:

$$Final\ Foverlty\ Score = \sum_{j=1}^{3} f_j S_j = f_1 S_1 + f_2 S_2 + f_3 S_3 = (0.5 \times 3) + (0.25 \times 7) + (0.25 \times 3) = 4.00$$

For *best novelty*, $S_1=10$ since the designer had created a stepping mechanism for propulsion, $S_2=7$ the same as in *final* novelty, $S_3=10$ as the designer had created reversed paddling brake. The *best* novelty score is:

$$Best\ Novelty\ Score = \sum_{j=1}^{3} f_j S_j = f_1 S_1 + f_2 S_2 + f_3 S_3 = (0.5 \times 10) + (0.25 \times 7) + (0.25 \times 10) = 9.25$$

3.2.3.2 Quantity evaluation. Let N_1, N_2, and N_3 be the count of ideas (leaf nodes) created and $N_{1\ \max}$, $N_{2\ \max}$, $N_{3\ \max}$, be the maximum N_1, N_2, and N_3 of all subjects. By inspecting genealogy trees for every subject, we identify $N_{1\ \max}=6$, $N_{1\ \max}=2$, and $N_{1\ \max}=4$ and from genealogy trees of this subject, $N_1=6$, $N_2=2$ and $N_3=4$. So the quantity subscores are:

$$Quantity\ Score\ for\ f_1, S_1 = 10 \times \frac{N_1}{N_{1\ \max}} = 10 \times \frac{6}{6} = 10$$

$$Quantity\ Score\ for\ f_2, S_2 = 10 \times \frac{N_2}{N_{2\ \max}} = 10 \times \frac{2}{2} = 10$$

$$Quantity\ Score\ for\ f_3, S_3 = 10 \times \frac{N_3}{N_{3\ \max}} = 10 \times \frac{4}{4} = 10$$

The overall quantity score is:

$$Overall\ Quantity\ Score = \sum_{j=1}^{3} f_j S_j = f_1 S_1 + f_2 S_2 + f_3 S_3 = (0.5 \times 10) + (0.25 \times 10) + (0.25 \times 10) = 10.00$$

3.2.3.3 Variety evaluation. Let S_{1k}, S_{2k}, and S_{3k} be predefined scores for level k in a genealogy tree of an idea set for Functions 1, 2, and 3 and b_k be a number of branches at level k in a genealogy tree. The variety subscore for each function is:

$$Variety\ score\ of\ f_1, S_1 = 10 \times \frac{\sum_{k=1}^{3} S_{1k} b_{1k}}{M_{f1\ \max}} = 10 \times \frac{(3 \times 6) + (5 \times 3) + (0 \times 1)}{6 \times 6} = 10 \times \frac{33}{36} = 9.12$$

$$Variety\ score\ of\ f_2, S_2 = 10 \times \frac{\sum_{k=1}^{3} S_{2k} b_{2k}}{M_{f2\ \max}} = 10 \times \frac{(0 \times 6) + (0 \times 3) + (2 \times 1)}{2 \times 6} = 10 \times \frac{2}{12} = 1.67$$

$$Variety\ score\ of\ f_3, S_3 = 10 \times \frac{\sum_{k=1}^{3} S_{3k} b_{3k}}{M_{f3\ \max}} = 10 \times \frac{(3 \times 6) + (2 \times 3) + (0 \times 1)}{4 \times 6} = 10 \times \frac{24}{24} = 10$$

The overall variety score is:

$$Overall\ variety\ score = \sum_{j=1}^{3} f_j S_j = f_1 S_1 + f_2 S_2 + f_3 S_3 = (0.5 \times 9.12) + (0.25 \times 1.67) + (0.25 \times 10) = 7.48$$

3.2.3.4 Quality evaluation. Let S_1, S_2, and S_3 be quality scores for Functions 1, 2, and 3. For Function 1, the subject applied paddling mechanism that accepts and converts paddling movement to the rotation of the wheels. The subject embedded needles to the wheels to increase friction. We rate it as feasible and good so the score of 10 is given for S_1. For Function 2, the design is a steering wheel that connected to steering column and a ski pad. We rate it as feasible and good so the score of 10 is given for S_2. For Function 3, the design is a wheel brake, which, although working, may not perform well on snow. We rate it as feasible but average so S_3 is given as 7. Therefore, the overall quality score is:

Table 2 Numbers of iterations

Subject	Global Iteration Loop PR	IS	CR	Total	Time (min)
1	10	8	13	31	36
2	6	4	6	16	12
3	2	1	3	6	20
4	5	3	7	15	20
5	5	4	6	15	11
6	6	5	8	19	23
7	3	4	6	13	20
8	5	2	7	14	12
9	3	1	3	7	13
10	10	5	12	27	26
11	7	3	9	19	24
12	5	3	9	17	32
13	2	2	3	7	11
14	3	1	3	7	16
15	4	2	6	12	26
16	6	3	7	16	24

Table 3 Frequency of iterations (loop per 10 min)

Subject	Global Iteration Loop PR	IS	CR	Total
1	2.78	2.22	3.61	8.61
2	5.00	3.33	5.00	13.33
3	1.00	0.50	1.50	3.00
4	2.50	1.50	3.50	7.50
5	4.55	3.64	5.45	13.64
6	2.61	2.17	3.48	8.26
7	1.50	2.00	3.00	6.50
8	4.17	1.67	5.83	11.67
9	2.31	0.77	2.31	5.38
10	3.85	1.92	4.62	10.38
11	2.92	1.25	3.75	7.92
12	1.56	0.94	2.81	5.31
13	1.82	1.82	2.73	6.36
14	1.88	0.63	1.88	4.38
15	1.54	0.77	2.31	4.62
16	2.50	1.25	2.92	6.67

$$\textit{Overall quality score} = \sum_{j=1}^{3} f_j S_j = f_1 S_1 + f_2 S_2 + f_3 S_3 = (0.5 \times 10) + (0.25 \times 10) + (0.25 \times 7) = 9.25$$

4 Results

4.1 Number and Frequency of Iterations. Numbers and frequencies of mental iteration of 16 subjects from the protocol study described in Sec. 3.1 are presented in Tables 2 and 3.

4.2 Design Scores. The evaluation was done by two assessors. The sketches of design concepts, both finished and unfinished, from the experiment were redrawn and presented to the assessors. Genealogy trees of idea generation were developed through the protocol studies. For variety and quantity, numbers of nodes and branches were obtained from the genealogy trees, then the scores were computed from the formula we described earlier. So they were independent to the assessor. For novelty, the score was also independent to the assessor as long as the same rate of novelty is used. For quality, although the rate of quality is predefined—i.e., 10 if feasible and good performance, 7 if feasible, 3 if infeasible—it still depends on the judgment of the assessor whether the idea is feasible and with good performance, feasible, or infeasible. The difference of quality scores rated between two assessors was less than 10% of the average value. Table 4 presents the scores from the averages of the two assessors. Table 5 summarizes final ideas generated by subjects for evaluating *final* novelty score and quality score. Table 6 summarizes the best-novel ideas for evaluating *best* novelty score.

4.3 Correlations. We now investigate the associations between design scores and number and frequency of iterations. *Pearson's correlation* coefficient (r) was used for preliminary testing of the association between numbers and frequencies of mental iterations in each loop and novelty, quality, variety, quantity scores. More testing may be needed to obtain accurate results.

Table 7 shows correlation coefficients of design scores and numbers of iteration. Table 8 shows correlation coefficients of design scores and frequencies of iterations. Significant level is chosen at 0.05 for a matter of convention.

For the numbers of iteration, as shown in Table 7, significant positive correlations exist between the numbers of every loop and design scores, except best and final novelty scores. For the best novelty, only positive correlation to IS loop is significant. For the final novelty, only negative correlation to PR loop is significant. For frequencies of iterations, there are significant negative correlations between the final novelty score and every loop, as shown in Table 8. But all other correlations are not significant.

Tables 7 and 8 also show correlations between each design score and correlations between iteration loops. There are significant positive correlations between quantity and best novelty, quantity, and variety. The positive correlation between quantity and novelty supports the argument that the more ideas, the better

Table 4 Design scores of all subjects

Subject	Best Novelty	Final Novelty	Variety	Quantity	Quality
1	9.25	6.00	7.48	10.00	8.25
2	7.50	5.00	5.00	3.54	4.00
3	6.50	5.75	5.00	3.75	2.25
4	6.50	3.00	3.75	4.79	3.00
5	6.50	6.50	2.22	5.63	1.50
6	5.75	7.00	2.92	4.79	10.00
7	8.50	6.00	5.00	4.38	4.75
8	4.75	5.25	6.25	4.79	8.00
9	5.00	4.50	1.67	2.92	2.25
10	6.00	4.50	6.25	5.63	7.50
11	6.00	4.50	1.67	3.54	7.50
12	7.50	5.25	5.42	5.83	8.00
13	7.75	8.00	2.50	3.54	6.25
14	6.50	3.75	0.83	2.92	2.25
15	5.75	5.50	0.83	3.54	5.50
16	7.50	3.75	5.00	4.38	9.00

Table 5 Summary of final ideas

Subject	*f*1 Propulsion	*f*2 Control Direction	*f*3 Brake
1	Paddle and 2 rear needle wheels	Steer with a front ski pad	Wheel brake with a braking stick
2	Sliding	Steer with ski pads	Cramping wheel brake
3	Lever mechanism and threads	Handle steer with small front wheels	Stop powering
4	Sketch board (no wheel)	Handle steer with a small front wheel	Stepping forks
5	Sketch board (with wheels)	Handle steer with a small front wheel	Cramping wheel brake
6	Paddle and 2 rear needle wheels	Handle steer with single front ski pad	Pivot and front ski brake
7	Paddle and threads	Handle steer with a front ski pad	Snow plug
8	Paddle and 2 rear needle wheels	Handle steer with 2 front wheels	Cramp brake
9	Paddle and (roller) threads	Handle steer with a front wheel	Cramping wheel brake
10	Paddle and 2 wheels	Handle steer with a front ski pad	Stop powering
11	Paddle and threads and gears	Handle steer with a front ski pad	Stop powering
12	Paddle and 2 rear wheels	Handle steer with a front wheel	Snow plug
13	Paddle and a thick needle wheel	Handle steer with a front ski pad	Clutch
14	Hand paddle and threads	Handle steer with a front wheel	Stop powering
15	Paddle and a thick needle wheel	Handle steer with a front ski pad	Side ski brake linkage
16	Paddle and 2 rear wheels	Handle steer with a front ski pad	Cramping wheel brake

Table 6 Summary of best-novel ideas

Subject	*f*1 Propulsion	*f*2 Control Direction	*f*3 Brake
1[a]	Stepping mechanism and rear needle wheels	Steer with single front ski pad	Reverse paddling
2[a]	Jetting with nozzle and ski pads	Steer with ski pads	Cramp brake
3	Lever mechanism and threads	Handle steer with 2 small front wheels	Not mentioned
4[a]	Sketch sliding board	Handle steer with 2 blades	Stepping forks
5[a]	Sketch board (with wheels)	Handle steer with a small front wheel	Cramp brake
6	Paddle and 2 rear needle wheels	Handle steer with single front ski pad	Pivot and front ski brake
7[a]	Pulling rope	Handle steer with single front ski pad	Snow plug (not feasible)
8[a]	Sketching blades	Handle steer with 2 blades	Cramp brake
9	Paddle and a (roller) thread	Handle steer with a front wheel	NA (assume cramp brake)
10[a]	Paddle and threads	Handle steer with a front ski pad	Stop paddling
11	Paddle and threads and gears	Handle steer with a front ski pad	Stop paddling
12[a]	Pulley system	Handle steer with a front wheel	Snow plug
13[a]	Paddle and thread	Handle steer with a front ski pad	Clutch
14	Hand paddle and threads	Handle steer with a front wheel	NA (assume stop powering)
15	Paddle and a thick wheel with a big stabilize board	Handle steer with a front ski pad	Side ski brake
16[a]	Lightweight snow shoes	Handle steer with a front ski pad	Cramp brake

[a]Denotes a subject whose best novelty score differs from the final novelty score

Table 7 Correlation matrix of design scores and numbers of iteration

	Best Novelty	Final Novelty	Variety	Quantity	Quality	PR Loop	IS Loop	CR Loop	Total Loops
Best Novelty	1.000								
Final Novelty	0.039	1.000							
Variety	0.427	−0.456	1.000						
Quantity	0.505[a]	−0.422	0.663[a]	1.000					
Quality	0.046	−0.242	0.340	0.465	1.000				
PR Loop	0.160	**−0.503**[a]	**0.512**[a]	**0.691**[a]	**0.643**[a]	1.000			
IS Loop	**0.532**[a]	−0.342	**0.553**[a]	**0.839**[a]	0.572[a]	0.806[a]	1.000		
CR Loop	0.249	−0.403	**0.587**[a]	**0.781**[a]	**0.725**[a]	0.935[a]	0.827[a]	1.000	
Total Loops	0.304	−0.440	**0.579**[a]	**0.801**[a]	**0.690**[a]	0.967[a]	0.904[a]	0.978[a]	1.000

[a]Denotes significant correlation at $\alpha=0.05$ (two tailed)

chance of novel ideas. For the correlations between iteration loops, all of them are positive and significant in both number and frequency.

The Pearson's correlation coefficient gives a quick glance at the relation between design scores and amount of iterations but it can mislead when there is data peculiarity and the correlation glosses over some major violation of the assumptions of Pearson's correlation coefficient. Such assumptions are a straight-line relationship and normally distributed characteristics [45]. To verify the result of a correlation analysis, one needs to inspect scatter plots in conjunction with the correlation matrix. An example of a scatter plot is presented in Fig. 7. We can see a fair negative correlation between *final novelty* and the number of PR loop but there is no correlation between *best novelty* and the number of PR loop.

5 Discussion

From the correlation analysis, the frequency of iterations did not correlate to any metric but the final novelty. On the other hand, the number of iterations was correlated to design metrics with an exception of novelty. The insignificant correlation between frequency of iterations and design metrics implies the tendency that it is not important whether designers iterate their thinking more or less frequently; it is the amount of iterations that matters. One can think that every design problem requires designers to perform a certain number of iterations to finish and achieve certain performance. In the following, we discuss the impact of iteration on each design metric.

5.1 Impact on Idea Novelty. As for the number of iteration, IS loop had positive impact on *best novelty* of ideas but it had no effect on the *final novelty* of the final design. This result implies that while more IS looping can lead to more novel ideas generated during the design process, there is no guarantee that the novel ideas will lead to the novelty of the final design. This implication is intriguing. Further examination of Table 7 shows that the PR loop had negative impact on the *final novelty* of the designed concept. Our interpretation of this intriguing result is that it is likely that novel ideas were created through iterations in the IS loop but later they were dismissed through iterations in the PR loop since revising the problem definition might have made the novel ideas obsolete. Generally speaking, novel ideas tend to be immature and hard to meet the evaluation criteria. Instead of exploring further on the novel ideas, designers tend to go back to redefine the problem and bring back "more familiar" solutions. As a result, many novel ideas created cannot reach the final design. The difference between best novel ideas and final ideas is illustrated in Table 5 and Table 6.

There are two possibilities that IS loop enables the creation of novel ideas. First, it is conceivable that IS loop facilitates stimuli of the design content that encourage designers to generate more new ideas. As a result, the chance of creating novel ideas increases through IS loop iterations. This interpretation is confirmed later when we discuss the positive correlation between the IS loop and the quantity score. The other possibility is that the stimuli facilitated by IS loop trigger designers, like a flashlight, to break through their mental block and create novel ideas. For example, during the design session, a subject stated, "…*I think it is acceptable. Ah, Snowboards!! Ah, you adapt sketch board to snow board. Maybe it is possible*…"

In contrast to the IS loop, there are two possibilities that the PR loop could reduce the novelty of the final design. First, it could prevent the creation of novel ideas. PR loop facilitates generation of new requirements and constraints. These new constraints could create a mental block or limit the design space that designers

Table 8 Correlation matrix of design scores and frequencies of iteration

	Best Novelty	Final Novelty	Variety	Quantity	Quality	PR Loop	IS Loop	CR Loop	Total Loops
Best Novelty	1.000								
Final Novelty	0.039	1.000							
Variety	0.427	−0.456	1.000						
Quantity	0.505	−0.422	0.663[a]	1.000					
Quality	0.046	−0.242	0.340	0.465	1.000				
PR Loop	−0.181	**−0.789**[a]	0.230	0.186	0.206	1.000			
IS Loop	0.281	**−0.557**[a]	0.239	0.351	0.171	0.806[a]	1.000		
CR Loop	−0.140	**−0.769**[a]	0.360	0.317	0.361	0.935[a]	0.827[a]	1.000	
Total Loops	−0.043	**−0.761**[a]	0.299	0.298	0.271	0.967[a]	0.904[a]	0.978[a]	1.000

[a]Denotes significant correlation at $\alpha=0.05$ (two tailed)

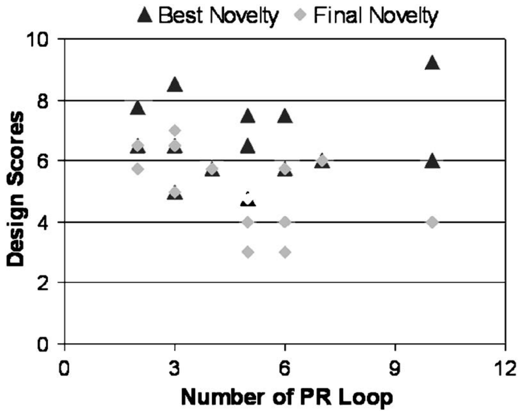

Fig. 7 An example of a scatter plot

would explore and consequently reduce the chance of creating novel ideas. Second, it dismisses novel ideas already created. When new requirements and constraints are introduced through the PR loop, previously created novel ideas may become obsolete. As a result, designers may leave these novel ideas for other familiar ones. Thus the novel ideas do not appear in the final design. For example, a subject stated that "*…but this small ski is very close to ski and self powered. But it not gonna fall in category of going uphill. You can't go uphill because it's really exhausting. It is really not improvement compared to ski. So that is kind of out of question and we won't consider that…*"

Besides the correlation to the number of iterations, the final novelty was also negatively correlated to the frequency of iterations of every loop. The strongest correlation occurred to PR loop ($r=-0.789$) followed by CR loop ($r=-0.763$) and IS loop ($r=-0.557$). This negative correlation can be interpreted as follows. Novel ideas often unexpectedly occur as a flash of designers' awareness. More frequent iteration implies less time designers spend in each looping. As novel ideas are discovered through the iteration, if the iteration occurs too quickly, these ideas can be easily lost due to the lack of incubation. As a result, the more frequent designers iterate, the less novelty in their final designs.

5.2 Impact on Idea Quantity. There was positive correlation between the quantity score and the number of iterations. The strongest correlation occurred to IS loop ($r=0.839$), while CR loop ($r=0.781$) and PR loop ($r=0.691$) had lower strength of correlation. On the other hand, the correlation to the frequency of iterations was not significant.

The positive correlation between the IS loop and the quantity score supports our early discussion that IS loop increases opportunity to generate more ideas by increasing the number of generated ideas. For the CR loop, it can be expected that concept reuse increases the chance to better use generated ideas and allows discarded ideas to be picked up and reused to compose other ideas. As a result, this loop increases quantity of ideas. Although PR loop has weaker correlation than other loops, its correlation strength is still considerably strong. Because PR loop is expected to facilitate information to redefine design problems, it increases more opportunity to decompose the problem in more different ways and generate different sets of requirements. These different sets of requirements consequently lead to more ideas.

5.3 Impact on Idea Variety. The result showed that variety was positively correlated to the number of iterations. The coefficients indicate approximately the same level of correlation for the three types of loop ($r=0.512$ for PR loop, $r=0.553$ for IS loop, $r=0.587$ for CR loop). However, the correlation to the frequency of iterations was not significant.

It can be expected that PR loop and IS loop increase variety of the ideas. The PR loop could increase variety of ideas by introducing different problem decompositions at higher abstraction level so that each idea has different principles leading to more variety. The IS loop could increase the variety by providing more stimuli to allow designers to create various ideas. On the other hand, the positive correlation between idea variety and CR loop counters our intuition. We expected that the CR loop reduces the variety of ideas because this loop allows the reuse of previously generated ideas, which should lead to more similarity and less variety. However, a careful examination of protocol data showed that the idea evolution occurred through iterations in every loop including CR loop. As more various ideas are generated, more CR looping occurred, too. So it should be said that CR loop is an indication, rather than driving force, of variety.

5.4 Impact on Idea Quality. The quality was positively correlated to the number of iterations, in which the CR loop has strongest correlation ($r=0.725$) followed by PR loop ($r=0.643$) and the IS loop ($r=0.572$). The correlation to the frequency of iterations was not significant.

Because the PR loop could lead to refining constraints and requirements, designers, through PR looping, could form a better set of requirements and constraints. This better-defined set of constraints and requirements guides designers to compose more better ideas. Designers can also improve the quality of ideas through iterations of CR loop because it increases the opportunity to reuse generated ideas and make them more refined and better. For the IS loop, because this loop increases the number of generated ideas, it can be expected that the chance of generating better ideas also increases.

5.5 Improving Mental Iteration. The empirical studies in this research provided evidence that mental iteration has a significant impact on design performance. As such, designers should be encouraged to have good mental iterations so that a better design performance can be achieved. One may suggest that designers iterate more IS loop but less PR loop if they want novel ideas, more IS loop and CR loop if they need as many designs as possible, more PR loop and CR loop if they prefer better design quality, or more every loop if they want more variety of ideas. Because most designers do not even consciously recognize these loops of mental iteration, methods or tools must be devised to encourage designers to perform desired mental iteration. It raises several questions for the future research—*how can we provide methodical and technological interventions to improve design iterations and design performance?* Many general methods, either conventional, intuitive, or discursive methods [12], have been developed to assist concept generation. It is conceivable that a direct or modified version of these methods or some new methods can be developed to encourage designers to have desired mental iterative behavior.

Moreover, positive correlation between design performance and the number of iterations raises the question whether better design performance can always be achieved by more iteration. One may hypothesize that there is a limit on the number of effective iterations. After the limit is exceeded, design performance is saturated. However, we inspected the scatter plots of design scores versus numbers of iteration to check a saturation pattern but there was no sign of such limit. It was possible that subjects in our study iterated much less than what could be.

Another issue worth mentioning is whether it is possible to emphasize one type of iteration loop over others. We have tested the correlation between each type of iteration loop and found strong positive correlations: $r=0.806$ for PR loop and IS loop, $r=0.935$ for PR loop and CR loop, and $r=0.827$ for IS loop and CR loop. This result suggests that each type of iteration loop does not occur without the others. Therefore, it may not be possible to have designers iterate one type of iteration loop without the others. Since there are both positive and negative correlations to number of iterations in different loops, a good compromise is required.

6 Concluding Remarks

In this paper, we performed an experimental study and correlation analysis to investigate relations between the measures of design performance and the number, frequency of mental iterations. The results suggested that (1) increasing number of iteration has positive impact on quality, variety, and quantity, but mixed effect on novelty, (2) the frequency of iteration does not seem to correlate to design performance except that it may have negative impact on the novelty of final design, and (3) increasing numbers of PR loop and IS loop may lead to decrease and increase of the novelty of design, respectively. We believe that the findings developed from this study would help advance the current understanding of mental iteration in terms of how mental iteration impacts designers' performance.

Our modeling and experimental study of mental iteration presented above has several limitations. First, the experiment included only 16 subjects. This could be a good starting point but experiments including more subjects need to be carried out before making broad generalizations from the results. Second, the subjects in our experiment were relatively inexperienced designers. There has been research on how experts perform design tasks differently from novice designers from different perspectives [38,46–49], including Adams [10] who investigated differences of iterative design behavior among freshmen and senior students. It could provide interesting insights for our research to involve experienced designers as subjects and study the difference of mental iteration between inexperienced and highly experienced designers. Finally, mental iteration was evaluated by the number, frequency, and percentage of each iteration loop. To provide more insight into mental iteration, other measures of mental iteration should be investigated. Such a measure may include the intensity, effectiveness, and efficiency of mental iteration. Our future work will address the above limitations.

References

[1] Costa, R., and Sobek II, D. K., 2003, "Iteration in Engineering Design: Inherent and Unavoidable or Product of Choice Made?" *Proceedings of ASME Design Engineering Technical Conferences*, Chicago, IL.

[2] Smith, R. P., and Eppinger, S. D., 1997, "Identifying Controlling Features of Engineering Design Iteration," Manage. Sci., **43**, pp. 276–293.

[3] Smith, R. P., and Eppinger, S. D., 1997, "A Predictive Model of Sequential Iteration in Engineering Design," Manage. Sci., **43**, pp. 1104–1120.

[4] Osborne, S. M., 1993, "Product Development Cycle Time Characterization through Modeling of Process Iteration," Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

[5] Eppinger, S. D., Nukala, M. V., and Whitney, D. E., 1997, "Generalized Models of Design Iteration using Signal Flow Graphs," Res. Eng. Des., **9**, pp. 112–123.

[6] Krishnan, V., Eppinger, S. D., and Whitney, D. E., 1997, "Simplifying Iterations in Cross Functional Design Decision Making," J. Mech. Des., **119**, pp. 485–493.

[7] Sobek II, D. K., Ward, A., and Liker, J. K., 1999, "Toyota's Principles of Set-Based Concurrent Engineering," Sloan Manage. Rev., **40**, pp. 67–82.

[8] Adams, R. S., and Atman, C. J., 1999, "Cognitive Processes in Iterative Design Behavior," *Proceedings of the 29th ASEE/IEEE Frontiers in Education Conference*, San Juan de Puerto Rico.

[9] Adams, R. S., and Atman, C. J., 2000, "Characterizing Engineering Student Design Process—An Illustration of Iteration," *Proceedings of the ASEE Annual Conference*, St. Louis, MO.

[10] Adams, R. S., 2001, "Cognitive Processes in Iterative Design Behavior," Ph.D. thesis, University of Washington, Seattle, WA.

[11] Adams, R. S., Turns, J., and Atman, C. J., 2003, "Educating Effective Engineering Designers: The Role of Reflective Practice," Des. Stud., **34**, pp. 275–294.

[12] Pahl, G., and Beitz, W., 1996, *Engineering Design: A Systematic Approach*, 2nd ed., Springer, New York.

[13] Ullman, D. G., 2002, *The Mechanical Design Process*, 3rd ed., McGraw–Hill, New York.

[14] Chusilp, P., and Jin, Y., 2004, "Cognitive Modeling of Iteration in Conceptual Design," *Proceedings of ASME Design Engineering Technical Conferences*, Salt Lake City, UT.

[15] Jin, Y., and Chusilp, P., 2005, "Modeling Mental Iteration in Conceptual Design," to appear in Des. Stud.

[16] Ullman, D. G., Dietterich, T. G., and Stauffer, L. A., 1988, "A Model of the Mechanical Design Process based on Empirical Data," Artif. Intell. Eng. Des. Anal. Manuf., **2**, pp. 33–52.

[17] Jansson, D. G., and Smith, S. M., 1991, "Design Fixation," Des. Stud., **12**, pp. 3–11.

[18] Schön, D., and Wiggins, G., 1992, "Kinds of Seeing and Their Functions in Designing," Des. Stud., **13**, pp. 135–156.

[19] Finke, R. A., Ward, T. B., and Smith, S. M., 1992, *Creative Cognition—Theory, Research, and Application*, MIT Press, Cambridge, MA.

[20] Maher, M. L., 1994, "Creative Design using a Genetic Algorithm," J. Comput. Civ. Eng., ASCE, pp. 2014–2021.

[21] Maher, M. L., Poon, J., and Boulanger, S., 1996, "Formalising Design Exploration as Co-evolution: A Combined Gene Approach," in Gero, J. S., and Sudweeks, F. (Eds.), *Advances in Formal Design Methods for CAD*, Chapman and Hall, London, UK.

[22] Shah, J., Vargas-Hernandez, N., Summers, J., and Kulkarni, S., 2001, "Collaborative Sketching as an Idea Generation Technique for Engineering Design," J. Creat. Behav., **35**, pp. 169–198.

[23] Cross, N., 2002, "Creative Cognition in Design: Processes of Exceptional Designers," in Hewett, T., and Kavanagh, T. (Eds.), *Creativity and Cognition*, ACM, New York.

[24] Benami, O., and Jin, Y., 2002, "Cognitive Stimulation in Conceptual Design," *Proceedings of ASME Design Engineering Technical Conferences*, Montreal, Canada.

[25] Benami, O., 2002, "Cognitive Approach to Creative Conceptual Design," Ph.D. thesis, University of Southern California, Los Angeles, CA.

[26] Mayer, R. J., 1992, "IDEF0 Function Modeling," *A Reconstruction of the Original Air Force Wright Aeronautical Laboratory Technical Report*, AFWAL-TR-81-4023 (the IDEF0 Yellow Book), Knowledge-based System Inc, College Station, TX.

[27] Gero, J. S., and McNeill, T., 1998, "An Approach to Analysis of Design Protocols," Des. Stud., **19**, pp. 22–61.

[28] Suwa, M., Purcell, T., and Gero, J. S., 1998, "Macroscopic Analysis of Design Processes Based on a Scheme for Coding Designers' Cognitive Actions," Des. Stud., **19**, pp. 455–483.

[29] Dorst, K., and Cross, N., 2001, "Creativity in the Design Process: Co-evolution of Problem-solution," Des. Stud., **22**, pp. 425–427.

[30] Dzbor, M., 1999, "Support for Problem Formalization in Engineering Design: An Enquiry into the Role of Knowledge-level Models," *The 10th International Symposium*, Vienna, Austria.

[31] French, M. J., 1985, *Conceptual Design for Engineers*, Design Council, London, UK.

[32] Cross, N., 2000, *Engineering Design Methods: Strategies for Products*, Wiley, New York.

[33] Kruger, C., and Cross, N., 2001, "Modelling Cognitive Strategies in Creative Design," *Computational and Cognitive Models of Creative Design V*, University of Sydney, Sydney, Australia.

[34] Van Someren, M. W., Barnard, Y. F., and Sandberg, J., 1994, *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes*, Academic, London, UK.

[35] Fricke, G., 1996, "Successful Individual Approaches in Engineering Design," Res. Eng. Des., **8**, pp. 151–165.

[36] VDI-Richtlinie 2225, 1977, *Technisch-wirtschaftliches Konstruieren*, VDI, Dusseldorf.

[37] Dorst, K., 1997, "Describing Design—A Comparison of Paradigms," Ph.D. thesis, Delft University of Technology, Netherlands.

[38] Atman, C. J., Chimka, J. R., Bursic, K. M., and Nachtmann, H. L., 1999, "A Comparison of Freshman and Senior Engineering Design Processes," Des. Stud., **20**, pp. 131–152.

[39] Dahl, D. W., Chattophadhyay, A., and Gorn, G. J., 2001, "The Importance of Visualization in Concept Design," Des. Stud., **22**, pp. 5–26.

[40] Besemer, S., and O'Quin, K., 1986, "Analyzing Creative Products: Refinement and Test of a Judging Instrument," J. Creat. Behav., **20**, pp. 115–126.

[41] Kardes, F. R., 1986, "Effects of Initial Product Judgments on Subsequent Memory-based Judgments," J. Consum. Res., **13**, pp. 1–11.

[42] Andrew, J., and Smith, D. C., 1996, "In Search of the Marketing Imagination: Factors Affecting the Creativity of Marketing Programs for Mature Products," J. Marketing Res., **33**, pp. 174–187.

[43] Shah, J. J., Vargas-Hernandez, N., and Smith, S. M., 2003, "Metrics for Measuring Ideation Effectiveness," Des. Stud., **24**, pp. 111–134.

[44] Song, S., and Agogino, A., 2004, "Insights on Designers' Sketching Activities in New Product Design Teams," *Design Engineering Technical Conferences*, Salt Lake City, UT.

[45] Levin, J., and Fox, J. A., 2000, *Elementary Statistics in Social Research*, 8th ed., Allyn and Bacon, Boston, MA.

[46] Waldron, M. B., Jelined, W., Owen, D., and Waldron, K. J., 1987, "A Study of Visual Recall Differences Between Experts and Naive Designer," in Eder, W. E. (Ed.), *Proceedings of the 1987 International Conference on Engineering Design*, Boston, pp. 86–93.

[47] Kavakli, M., and Gero, J. S., 2001, "Strategic Knowledge Differences Between an Expert and a Novice," in Gero, J. S., and Hori, K. (Eds.), *Strategic Knowledge and Concept Formation III*. Key Centre of Design Computing and Cognition, University of Sydney, Australia, pp. 55–68.

[48] Lewis, W. P., and Bonollo, E., 2002, "An Analysis of Professional Skills in Design Implications for Education and Research," Des. Stud., **23**, pp. 385–406.

[49] Ahmed, S., Wallace, K. M., and Blessing, L., 2003, "Understanding the Differences between How Novice and Experienced Designers Approach Design Tasks," Res. Eng. Des., **14**, pp. 1–11.