Artificial Intelligence for Engineering Design, Analysis and Manufacturing (2010), *24*, 161–178.

doi:10.1017/S089006041000003X

Investigating novelty–outcome relationships in engineering design

V. SRINIVASAN AND AMARESH CHAKRABARTI

Centre for Product Design and Manufacturing, Indian Institute of Science, Bangalore, India

(RECEIVED April 22, 2009; ACCEPTED October 16, 2009)

Abstract

Design creativity involves developing novel and useful solutions to design problems. The research in this article is an attempt to understand how novelty of a design resulting from a design process is related to the kind of outcomes, described here as constructs, involved in the design process. A model of causality, the SAPPhIRE model, is used as the basis of the analysis. The analysis is based on previous research that shows that designing involves development and exploration of the seven basic constructs of the SAPPhIRE model that constitute the causal connection between the various levels of abstraction at which a design can be described. The constructs are state change, action, parts, phenomenon, input, organs, and effect. The following two questions are asked. Is there a relationship between novelty and the constructs? If there is a relationship, what is the degree of this relationship? A hypothesis is developed to answer the questions: an increase in the number and variety of ideas explored while designing should enhance the variety of concept space, leading to an increase in the novelty of the concept space. Eight existing observational studies of designing sessions are used to empirically validate the hypothesis. Each designing session involves an individual designer, experienced or novice, solving a design problem by producing concepts and following a think-aloud protocol. The results indicate dependence of novelty of concept space on variety of concept space and dependence of variety of concept space on variety of idea space, thereby validating the hypothesis. The results also reveal a strong correlation between novelty and the constructs; correlation value decreases as the abstraction level of the constructs reduces, signifying the importance of using constructs at higher abstraction levels for enhancing novelty.

Keywords: Concept; Idea; Novelty; Novelty Assessment Method; Variety; Variety Assessment Method

1. INTRODUCTION

The current market scenario is competitive, and this demands a company to retain or improve its competitive edge, to have a wider customer reach. There are several ways in which this could be achieved; one of them is by developing creative products. Engineering design is a process that spans from the identification of a need to a stage, where information with sufficient detail about an engineering solution can be developed, such that the solution can be produced and implemented to fulfill the need. Chakrabarti (2002) considered engineering design as a central part of the product development process, and it is distinguished from other aspects of engineering by its creative aspects, whereby novel products are conceived. Therefore, by helping a company with an approach to support creativity, especially in the designing phases of product development, should help the company have a wider customer reach.

Reprint requests to: Amaresh Chakrabarti, Centre for Product Design and Manufacturing, Indian Institute of Science, Bangalore 560 012, India. E-mail: ac123@cpdm.iisc.ernet.in

A variety of definitions have been proposed for creativity. After analyzing a comprehensive set of definitions and integrating the generic, encompassing elements from these, Sarkar (2007) proposed a common definition for creativity in engineering design: "Creativity in engineering design occurs through a process by which an agent uses its ability to generate ideas, solutions or products that are novel and useful."

Designing involves multiple facets: people, product, process, tools, organization, and environment in which designing takes place (Blessing et al., 1995). The characteristics of a designed product are influenced by some or all of the above facets. For example, creativity of a product depends on the creativity of its designers (people); the underlying physical phenomena in the product (product); the use of design methods like brainstorming, the Innovation Situation Questionnaire, and the theory of inventive problem solving (process); and the use of tools like a sketching tool (tools). Designing is also characterized by interactions, within and among the facets. Consequently, the product characteristics become a complex function of the facets.

In this article, we study the relationship between novelty (an essential component of creativity in engineering design) of a product and outcomes, which are the product facets that are used in designing the product. Such a study becomes relevant because it provides an understanding for answering the following questions:

1. How is a novel product created?
2. How can we improve the chances of designing novel products?

This article is categorized into the following sections: Section 2 reports relevant findings from literature and identifies the research questions, Section 3 proposes a hypothesis that encapsulates a possible relationship between novelty of a product and the constructs used in designing the product, Section 4 explains the research methodology adopted to tackle the research objectives, Section 5 reveals the results from an empirical validation of the hypothesis, Section 6 discusses the contributions of this article, and Section 7 summarizes the research in this article and charts out directions for further research.

2. LITERATURE SURVEY

This section reports relevant findings from the literature.

2.1. Novelty

Novel means "new and original, not like anything seen before" and novelty is the "quality of being new and unusual and something that has not been experienced before, and so is interesting" (Cambridge University Press, 2009). Novelty resembles something not formerly known (Sternberg & Lubart, 1999). Novelty resembles unusualness or unexpectedness (Shah et al., 2003). According to Sarkar (2007), novelty happens when an agent generates an outcome without replicating any existing outcome(s). Shah et al. (2003) and Lopez-Mesa and Vidal (2006) used "infrequency" as a measure of novelty. "Nonobviousness" is used as a measure while assessing novelty in patent documents (Franzosi, 2006). Shah et al. (2003), Lopez-Mesa and Vidal (2006), and Sarkar (2007) considered novelty as one of the measures of creativity of engineering products.

2.2. Types of novelty

Boden (1999) classified creativity into one of the following types: psychological creativity (P-creativity) and historical creativity (H-creativity). P-creativity is defined with reference to the ideas produced by an individual, whereas H-creativity is defined with reference to all the ideas produced during the whole of human history. H-creativity includes P-creativity because a historically creative idea is new to all, including the individual. Because novelty is a measure of creativity and is included within creativity it can be argued that the existence of P-/H-novelty on similar lines should also be valid. An ideology similar to Boden (1999) in the classification is followed by Shah et al. (2003), who classified novelty into three different levels: personal novelty, societal novelty, and historical novelty. In personal novelty, an individual discovers or creates products or ideas that are new according to that individual. In societal novelty, a product or idea is new to all people in a particular society, regardless of whether the product or idea is commonplace in other societies. In historical novelty, a product or idea is the first of its kind in the history of all societies and civilizations. Historical novelty subsumes personal and societal novelty, and societal novelty encompasses personal novelty. In this research we focus on historical novelty because it includes other types of novelty.

2.3. Importance of novelty

Westwood and Sekine (1988) argued that creative acts and innovative processes are the essential ingredients of any competitive industrial enterprise. They also mentioned that although marketing largely determines where the enterprise shall go, and management, when and in what style, creativity and innovation are the crucial factors that determine whether the enterprise will arrive at its preferred destination in a timely and profitable manner or not. Ottosson (1995) felt that creative products might be used to increase the price of products, and hence get a larger market share. Zimmerman and Hart (1998) felt that without creative problem solving, products will be traditional, without a creative edge, which can cause losses at the market place. According to Eder (1995), creative solutions can cause innovation in all design areas. Shalley (1991) and Unsworth (2001) suggested that some level of creativity is required almost in any job. This shows the importance of creativity, and hence that of novelty, which is one of the measures of creativity. Thring and Laithwaithe (1977) argued that society needs more inventions than ever before as the world's resources become scarce while one-third of the world's rapidly growing population is undernourished. It can be argued that all inventions are associated with some degree of originality and therefore novelty. In addition, increasing competition in the world market has forced companies to look for new ideas to improve the quality of products (Molina et al., 1995); this further signifies the importance of novelty.

Creativity assessment should also help assess the amount of innovation taking place in design firms and identify better inventors or designers. A method that can help identify creativeness of engineering products is required. According to Sarkar (2007), measuring novelty is important because it helps determine a design's newness and patentability, serves as a criterion for comparing designer's capability, and ascertains the potential market of a product. Novelty is a primary measure of creativity, whose measurement is useful for research, team recruitment, and so forth (Lopez-Mesa & Vidal, 2006).

2.4. Conceptual design

Pahl and Beitz (1996) defined conceptual design as a phase of designing where principles of solutions are developed.

French (1999) distinguished conceptual design from the other phases in terms of its greater fluidity and flexibility it offers. According to French (1999), this phase offers maximum scope for most striking improvements. Pahl and Beitz (1996) felt that a lasting and successful solution is more likely to spring from the choice of the most appropriate principles than from exaggerated concentration on technical details. Being an early phase, any changes made in this phase are less costly and more effective, and also determine the nature of the solution to be developed later in the design process. Berliner and Brimson (1988) pointed out that, on average, about 80% of the cost of a product over its total life cycle is committed during the conceptual design phase.

2.5. Physical laws and effects

Physical laws and effects are principles of nature that govern a change (Chakrabarti et al., 2005). A physical law in its widest sense represents the functional connection between variables, geometrical parameters, material constants, and basic constants (Zavbi & Duhovnik, 2000). Natural laws, which comprise physical laws and effects, provide important information for supporting invention and development of artifacts (Koyama et al., 1996). Koyama et al. (1996) developed a four-step model of the invention process of creative engineers:

1. Convert the functional requirements into the behavior of an artifact.
2. Retrieve information that may be helpful to predict the behavior of the artifact.
3. Assume proper arrangement of the parts to implement the artifact.
4. Apply the information to the resulted artifact to prove that the artifact can really achieve the required behavior.

Koyama et al. (1996) claimed that natural laws are the most important information for Steps 1 and 4.

Zavbi and Duhovnik (2000) argued that if operation of existing technical systems can be explained using physical laws, then these can also be used to design such systems. They consider physical laws as the basic and richest source for designing; basic, because no technical system operates contrary to them, and richest, because each physical law can be materialized in several topologies, each topology in several forms and each form in several materials. Designing at the level of physical laws also prevents a designer's fixation on adaptations of the existing solutions or composition of solutions from the existing components (Zavbi & Duhovnik, 2001).

A conceptual solution can be described as a causal network of physical effects (Chakrabarti et al., 1997). Conceptual design at the level of physical laws enables greater ability to innovate (Zavbi & Duhovnik, 2001). Murakoshi and Taura (1998) pointed out that laws and effects help synthesize novel artifacts. However, this has not been verified empirically. Murakoshi and Taura (1998) argued that synthesizing artifacts directly from physical effects is hard because effects have been created and described by scientists primarily for explanation of phenomena rather than for synthesizing artifacts that embody these phenomena; synthesis using laws and effects requires more than a straightforward application. Therefore, in the current form, effect representations are ill equipped in aiding synthesis in a substantial way. Srinivasan and Chakrbarti (2008) observed that designers do not use adequate laws and effects in designing. Their observations are based on six different observational studies of designing sessions, where designers in teams solve a design problem by producing conceptual solutions through a think with discuss-aloud protocol. Sarkar and Chakrabarti (2007*a*) also reported similar observations using a different set of observational studies; their research objective was to understand the different search spaces explored by designers while solving a problem.

2.6. SAPPhIRE model of causality

The SAPPhIRE model of causality was developed by Chakrabarti et al. (2005; Fig. 1) to explain the causality of natural and engineered systems. The model gets its name from the highlighted letters of its constructs: state change, action, parts, phenomenon, input, organs, and effect. Note that the term effect is used to collectively mean a combination of physical laws and effects. The constructs have been integrated from different approaches in the literature to create a more comprehensive model of causality: Umeda's function–behavior–state model (Umeda et al., 1996), Hubka's theory of technical systems (Hubka, 1976), Andreasen's domain theory (Andreasen, 1980), and Yoshioka and Tomiyama's concept of the Metamodel (Yoshioka & Tomiyama, 1997). The definition of the constructs have been rephrased in Srinivasan and Chakrabarti (2009; Table 1) to provide greater clarity in understanding the constructs and hence usage of the model. Chakrabarti et al. (2005) stated that a description of functionality can take different forms: an action description (e.g., cool

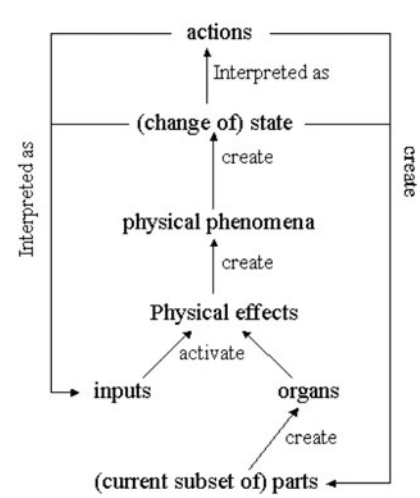

Fig. 1. The SAPPhIRE model of causality according to Chakrabarti et al. (2005).

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

Table 1. *Rephrased definitions of SAPPhIRE constructs*

Construct	Definition
Phenomenon	An interaction between a system and its environment
State change	A property at an instant of time of a system (and environment), which is involved in an interaction between the system and the environment; as a consequence of an interaction, a property of the system (and environment) changes and this is referred to as a state change.
Effect	Principle of the nature that underlies or governs an interaction
Action	An abstract description or high-level interpretation of an interaction between a system and its environment
Input	A physical quantity or variable that comes from outside the system boundary that is essential for an interaction between a system and its environment; this quantity can take the form of material, energy, or information.
Organs	The properties and conditions of a system and its environment required for an interaction between them
Parts	A set of physical components and interfaces that constitute the system and its environment

The definitions are according to Srinivasan and Chakrabarti (2009).

body, move body, generate current, etc.), input–output of a system (e.g., temperature difference as input to heat transfer as output, acceleration as input to displacement as output, potential difference as input to current as output), and state change (e.g., change in temperature, change in spatial location, change in current). The ability of the model to accommodate functionality in its different forms and link them together provides a greater richness in the description of functionality. Physical phenomena and physical laws/effects together are rarely supported by a single model or approach in the literature. The use of these constructs together and their links with functionality provides a richer description of behavior. Action, state change, and input form the higher levels of abstraction. Physical phenomenon and effect comprise the intermediate levels of abstraction. Organs and parts form the lower levels of abstraction. A model of causality with the refined definitions is explained as follows in Srinivasan and Chakrabarti (2009): a set of components and interfaces that constitutes a system and its environment (parts) creates a set of properties and conditions of the system and its environment (organs). When the system and its environment are not in equilibrium, there is a transfer of a physical quantity in the form of material, energy, or signal (input) across the system boundary. This physical quantity, in combination with a particular set of properties and conditions (organs), together activate a principle (effect). Activation of this principle creates an interaction between the system and its environment (phenomenon). The interaction between the system and its environment creates a change in property of the system (state change). The change in property can be interpreted at a higher level of abstraction (action).

An example (Fig. 2 and Fig. 3) is shown for a better understanding of the model. Let us assume that a body is kept in an air medium (parts). The relevant properties and conditions that

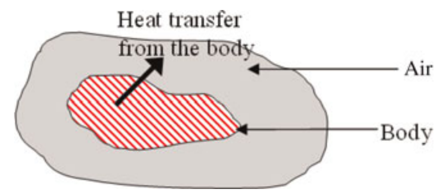

Fig. 2. An example of how a hot body cools down. [A color version of this figure can be viewed online at journals.cambridge.org/aie]

can be created from the parts are the fluidic property of the medium, the value of the surface area of the body (A), and the heat transfer coefficient between the body and the surrounding medium (h), which is a function of the geometry of the body and the nature of the surrounding fluid medium (organs). Let us assume that the body is at a higher temperature than the medium ($T_b > T_s$), resulting in a temperature difference between the body and the medium ($T_b > T_s$) (input). The input and the organs activate the convection heat transfer law [$Q = hA(T_b - T_s)$] (effect). This effect creates a heat transfer from the body to the surrounding medium (phenomenon). The phenomenon creates a decrease in the temperature of the body (state change). This state change can be interpreted as cooling of the body (action). Srinivasan and Chakrabarti (2009) demonstrated the capabilities of the model for analysis and synthesis through examples drawn from multiple domains and drew the following conclusions from these demonstrations:

1. The model can support analysis and synthesis.
2. The model can support analysis and synthesis of multidisciplinary systems.

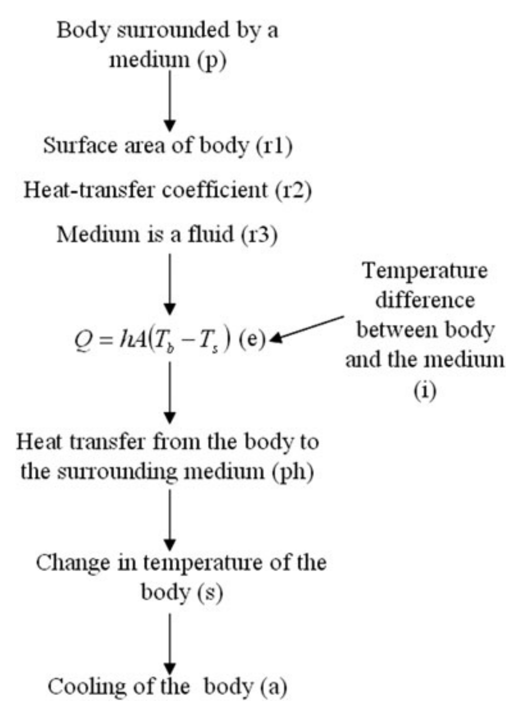

Fig. 3. An example demonstrating the use of the SAPPhIRE model.

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

3. The model can support analysis and synthesis of simple and complex systems.
4. In analysis and synthesis, the model requires and develops information, respectively, about a system and its environment.
5. The model supports conceptual and early embodiment phases of designing.

Designing is often seen as a combination of synthesis and analysis (Srinivasan & Chakrabarti, 2008). Because the SAPPhIRE model can help carry out both synthesis and analysis, it can be argued that it should also be able to help in designing. One way of validating this is by checking if the constructs of SAPPhIRE are already present in designing where designers are not prescribed to use any particular model of designing. Srinivasan and Chakrabarti (2008) evaluated the constructs of SAPPhIRE to check if these can be used in designing. This is achieved by checking whether or not the constructs are present in a series of designing sessions; these designing sessions were undertaken much before the model was developed. Six existing observational studies of designing sessions from an earlier research in Chakrabarti (2003), involving designers in teams of three, solving design problems by producing conceptual solutions, are used for this purpose. The designers in these sessions were instructed to follow one of the following methods: functional analysis, innovation situation questionnaire, and ideal design approach, but not enforced to follow any particular methodology, that is, they were allowed to work in their natural way. A two-way validation is performed as follows:

1. Check if all instances in the observational sessions could be represented by the SAPPhIRE constructs.
2. Check if all SAPPhIRE constructs have instances in the observational sessions.

Observations from the six (1–6) studies revealed that all the constructs of the SAPPhIRE model are present in designing. However, the constructs are not present in the expected numbers (Fig. 4). A large number of instances of action-, input-, and part-level descriptions are used by designers (note that descriptions of input level are considered under action level). A relatively small number of instances of phenomenon, effect, and organ level are observed. Similar findings were also reported by Sarkar and Chakrabarti (2007*a*); they analyzed the different search spaces explored by designers, using a different set of observational sessions, each involving an individual designer, experienced or novice, solving a conceptual problem. One would normally expect the number of instances to increase from a higher level of abstraction to a lower level of abstraction, that is, from action to parts (as indicated by the dashed black line in the graph). This is because every action can be satisfied by multiple alternative state changes, every state change can be satisfied by multiple alternative phenomena, and so forth, culminating in a diverging treelike structure. This feature of the model has also been demonstrated by Srinivasan and Chakrabarti (2009) using examples taken from multiple domains. As the graph shows, there is a considerable dip in the phenomenon, effect, and organ levels. This could be for several reasons: the designers observed did not possess a good knowledge of phenomena and effects; the designers did not know to how use laws and effects in designing; and in the sessions, the designers were not instructed to use laws and effects in designing. Similar observations from the two different sets of observational studies (Sarkar & Chakrabarti, 2007*a*, 2008) revealed that designers are not equally well versed with all of the constructs of SAPPhIRE. This becomes a significant weakness, especially if laws and effects have an influence on product novelty.

2.7. Novelty assessment methods

Amabile (1996) suggests the use of experts to identify what is "creative," because ultimately for any measure of creativity to be valid the results should match the notion held by experts. Sarkar (2007) feels that identification of a novel product in an absolute sense is difficult, because it is difficult to be aware of all the products available in all countries. He proposes that an ideal resource containing information of all products from

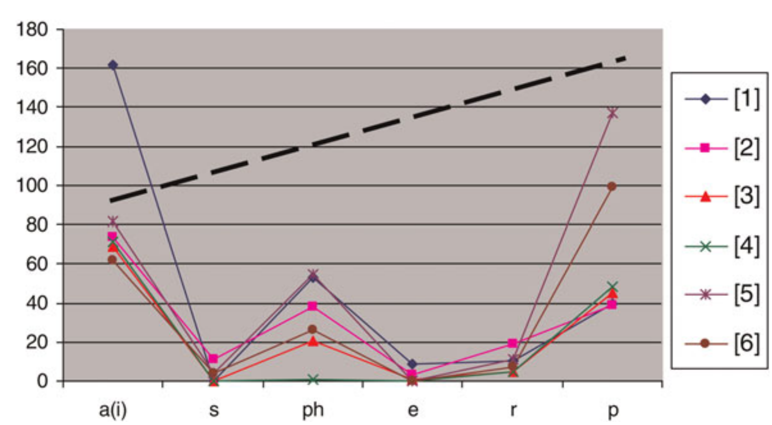

Fig. 4. The frequency of SAPPhIRE constructs from Srinivasan and Chakrabarti (2008). [A color version of this figure can be viewed online at journals.cambridge.org/aie]

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

all domains and their characteristics (e.g., an extensive searchable database) could be a solution to solve this issue. However, in the absence of such an information base, he believes that novelty assessment could be done by experienced designers who have knowledge of the domain(s) of the product whose novelty is to be assessed. He also adds that, generally in a design firm, creativity of conceptual solutions is judged by experienced designers who also make a decision on which of the concepts should be developed further into products. Even in patent offices, novelty and usefulness of products (measures of creativity) are judged by experienced designers who have knowledge in the same or related domains. Sarkar (2007) added strength to his arguments by pointing out that in novelty assessment, product users will be less reliable than experienced designers, because it is unlikely that they will have the background to understand the degree of novelty that is domain dependent. In contrast, designers with experience in related domains will be better aware of products in these domains, thereby providing more reliable judgment about novelty.

A number of researchers have proposed methods to help assess and determine novelty of products. The work of Saunders (2002) addressed finding novelty of patterns restricting primarily to aesthetics, through these questions: how often have similar patterns been experienced, how similar have these patterns been, and how recently have the patterns been experienced. Novelty is detected computationally by using processes that estimate these properties for a given stimulus pattern and a representation of previous stimuli. Saunders measured novelty of a pattern in terms of the frequency of similar patterns, similarity of patterns, and similarity in terms of time.

Shah et al. (2003) proposed two approaches for measuring novelty. In the first approach, all the ideas that are nonnovel (expected, usual) are collected before analyzing any data. In the second approach, all the ideas produced by the designers are collected. All of the key attributes for a design task and also different ways by each attribute can be accomplished, are identified. The number of instances of each solution method in the entire collection of ideas is counted: the fewer there are, the more novel the solution is. The measurement procedure for novelty is as follows: the problem is first decomposed into its key functions or characteristics. Every idea produced is analyzed by first identifying what functions it satisfies and describing how it fulfills these functions at the conceptual and embodiment level. Each description is then graded for novelty according to one of two approaches. It is possible to compute a total score for novelty for each idea by applying the weights to each function and stage. The overall novelty of each idea can be computed from the following:

$$M_1 = \sum_{j=1}^{m} f_j \sum_{k=1}^{n} s_{1jk} p_k, \tag{1}$$

where M_1 is the overall novelty score for an idea with m functions or attributes and n stages. Weights (f_j) are assigned according to the importance of each function or characteristic to compute an overall score. Further, each function may be addressed at the conceptual and embodiment stage and weights (p_k) assigned according to the stage's importance. The calculation of s_{1jk} depends on the approach chosen. For the first approach (a priori knowledge) a universe of ideas for comparison is subjectively defined for each function or attribute, and at each stage. A novelty score s_{1jk} is assigned to each idea in this universe. To evaluate the function and stage of an idea, a closest match is found. For the second approach, s_{1jk} is calculated from the following:

$$s_{1jk} = \frac{T_{jk} - C_{jk}}{T_{jk}} \times 10, \tag{2}$$

where T_{jk} is the total number of ideas produced for function (or key attribute) j and stage k, and C_{jk} is the count of the current solution for that function (or key attribute) and stage. Multiplying by 10 normalizes the expression. Apart from measuring novelty, the authors also measure variety by examining how each function is satisfied. A variety rating of a group of ideas is based on how different two ideas are from each other. For instance, two ideas are very different if they use different physical principles to satisfy the same function and are only slightly different if they differ only in some construction level detail. The variety is calculated from

$$M_3 = \sum_{j=1}^{m} f_j \sum_{k=1}^{4} s_k b_k / n, \tag{3}$$

where M_3 is the variety score, b_k is the number of branches at level k, m is the total number of functions, and s_k is the score for level k (four scores 10, 6, 3, and 1 are assigned for physical principle, working principle, embodiment, and detail levels, respectively).

Chakrabarti and Khadilkar (2003) developed a method based on the following rules: novelty of a product cannot be assessed without assessing its similarity or difference with existing products as reference, and several levels of novelty exist because of differences at principle, technology, and implementation levels. The criteria, based on which novelty can be assessed, have two levels: vertical and horizontal. Vertical-level criteria are fundamental product characteristics and comprise need, task, subsystem structure, working principle, technology, and implementation. Horizontal-level criteria are based on the relative importance in the overall functioning at that level and comprise main function, supplementary function(s), and additional function(s). A given product is compared with a reference product, and all the differences in horizontal and vertical criteria are identified in terms of weightage assigned to each criterion. The novelty value of each difference is computed by multiplying the weightage at that vertical level and horizontal level. This product is multiplied with the horizontal-level weightage at one level above in the vertical direction, and multiplication continues until the highest level in vertical direction is reached. The novelty value of each difference is summed up to get the novelty value of the product.

Lopez-Mesa and Vidal (2006) developed two methods for measuring novelty of design alternatives. Each design alternative is classified in terms of its action function, structure, and detail. In the first method, novelty is measured in terms of newness with respect to the current paradigm, where each design alternative is classified into one of the four change type patterns (in ascending order of measure of newness): type 1, type 2, type 3, and type 4. By counting the relative number of alternatives under each type belonging to each designer or design team involved, novelty of the designer or design team can be assessed. In the second method, novelty is measured in terms of nonobviousness of the outcomes. Nonobvious solutions are those that are produced by few individuals/teams, and therefore, the fewer the number of such solutions produced across the individuals or teams concerned, the greater is the nonobviousness. Nonobviousness is measured at different levels, where the number of levels is equal to the number of subjects (team/individual) being compared. For instance, if there are four teams, then the levels are: solutions produced by four teams, solutions produced by three teams, solutions produced by two teams and solutions produced by one team.

Sarkar and Chakrabarti (2007*b*, 2008) developed a method (Fig. 5) for assessing qualitative relative degree of novelty of

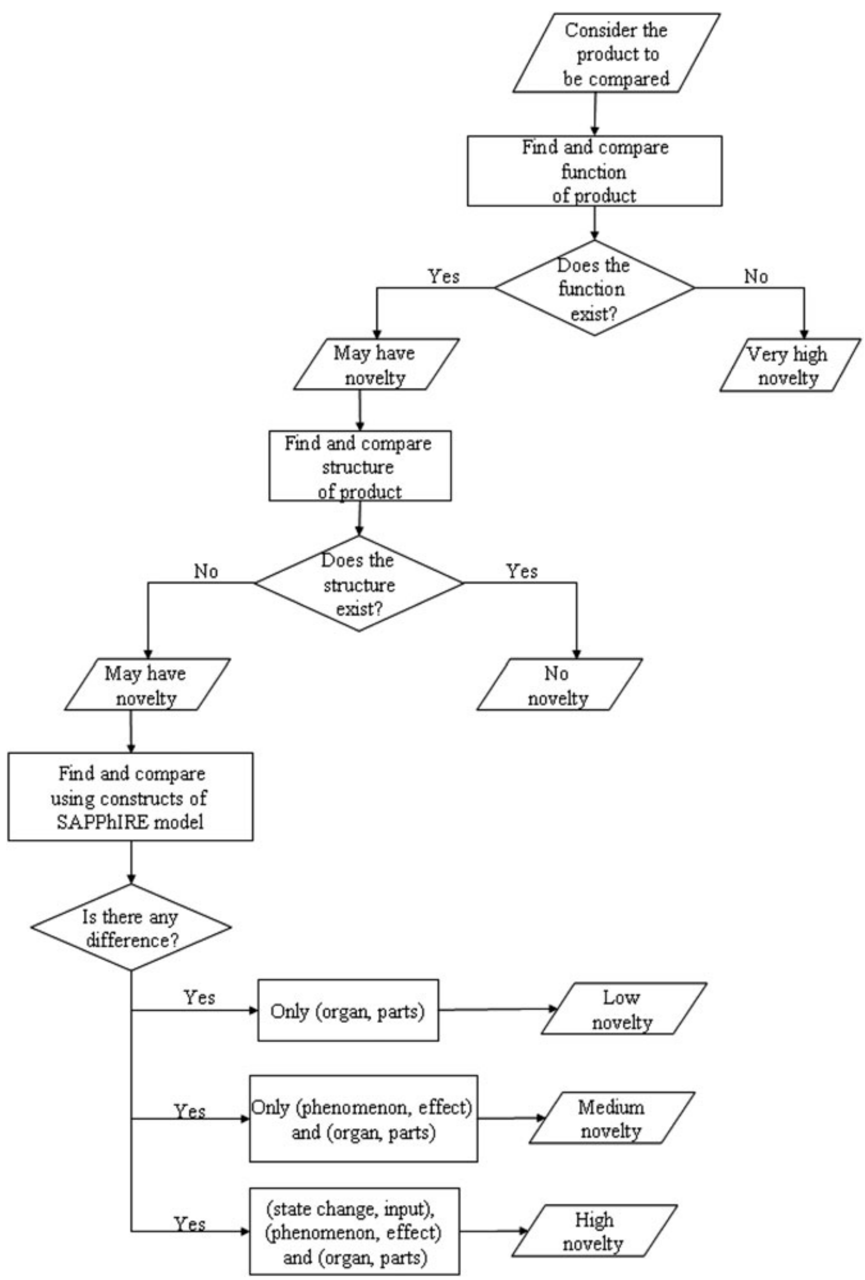

Fig. 5. The novelty assessment method of Sarkar and Chakrbarti (2007*b*, 2008).

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

an engineering product using the constructs of the product. The various degrees of novelty are: very high, high, medium, or low. According to the authors, the method employs function–behavior–structure (FBS) and SAPPhIRE models together; the FBS model is used first for determining novelty and then SAPPhIRE model is used to assess the relative degree of novelty. The method is based on the differences between the product (whose novelty has to be measured) and existing products, in terms of the SAPPhIRE constructs at different abstraction levels. If a product whose novelty has to be assessed is different from existing products at a higher level of abstraction, then it has a higher novelty; novelty decreases as the difference reduces to lower levels of abstraction. If the product is not different from the existing products, the product is not novel. This method does not use all the constructs of both the models (behavior of FBS and action of SAPPhIRE). It can be argued that the method need not use both FBS and SAPPhIRE because in a richer sense action, state change, and input comprise function, while phenomenon, effect, and organs comprise behavior, and organs and parts comprise structure. While comparing using SAPPhIRE constructs, organs and parts, phenomenon and effect, state change and input can be treated separately. This is so because the same organ, phenomenon, and state change can be accomplished by alternative parts, effects, and phenomena, respectively.

2.8. Summary

This section summarizes the findings from literature (Sections 2.1–2.7) and relates these findings to one another to frame the research questions. Novelty encompasses both newness and originality. It is important for several reasons (Section 2.2). A number of methods for assessing novelty of products exist in the literature (Section 2.7). Conceptual design being an early stage can accommodate changes that are most striking and effective and yet less expensive (Section 2.4). Novelty in designs can be addressed by the use of laws and effects in designing. However, no empirical studies were found that corroborate this relationship (Section 2.5). The SAPPhIRE model of causality uses its constructs to explain the causality of natural and engineered systems. Although the model can assist in analysis and synthesis, it has not been tested for its capabilities in assisting designing. Preliminary investigations using protocol studies revealed that the constructs of the SAPPhIRE model are present in of designing; the studies have been confined primarily to conceptual and early embodiment phases. However, the constructs are not present in the expected numbers; very few phenomena, effects, and organs are reported, which is against what would be logically expected. This might seriously hinder the chances of enhancing novelty, if it is strongly related to physical laws and effects. Under these circumstances, it is important to empirically check the relationship between novelty and physical laws and effects, as well as check the relationships between novelty and the other constructs of SAPPhIRE. The following two questions are asked:

1. Is there a relationship between novelty and the constructs of SAPPhIRE model?
2. If there is a relationship, what is the degree of this relationship?

3. UNDERLYING RESEARCH HYPOTHESIS

This section is divided into two parts. Section 3.1 defines the different terms involved in the hypothesis, and Section 3.2 provides an explanation of the hypothesis.

3.1. Terminology

Before explaining the hypothesis for this study, we define the terms involved (Fig. 6): a *concept* is defined here as an entity that satisfies an overall function. A concept is a solution that satisfies most of the requirements identified for a problem. In Figure 6, C1, C2, C3, and C4 are different concepts that

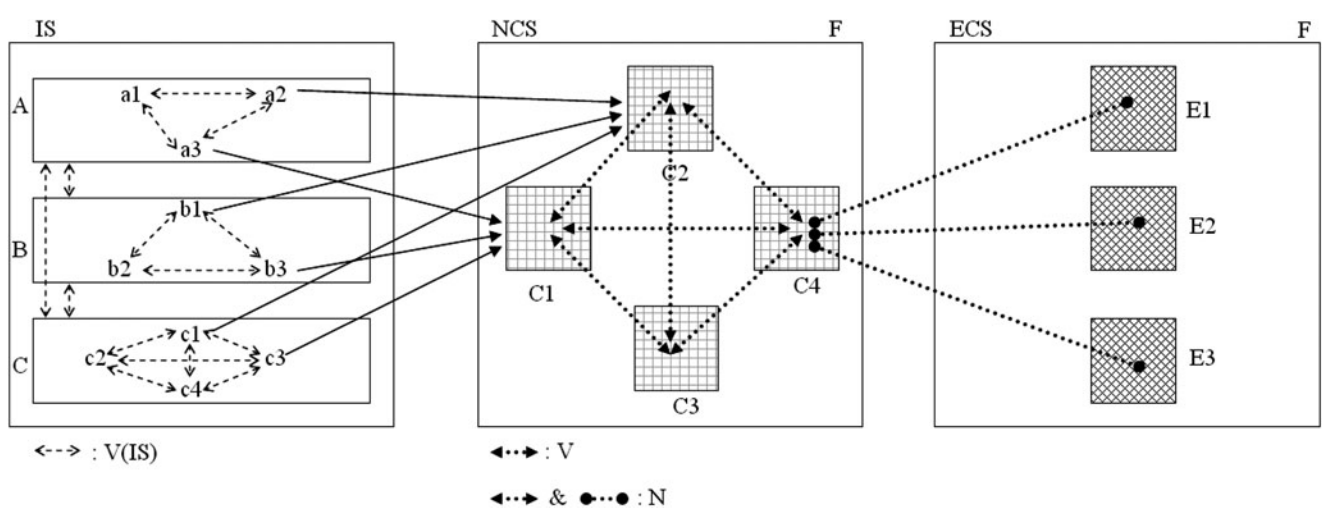

Fig. 6. A pictorial representation of idea space (IS), new concept space (NCS), and existing concept space (ECS).

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

satisfy an overall function *F*. For example, a broom and a vacuum cleaner are two alternative concepts that are used to clean dust, which is a function.

An *idea* is defined here as an entity at a particular level of abstraction. An idea is a solution that satisfies requirements at a particular level of abstraction only. An idea is a constituent of a concept. For instance, ideas a3, b3, and c3, constitute concept C1, as shown in Figure 6. For example, suction is an idea at the phenomenon level, a broom handle is an idea at the part level, and so forth.

An *idea group* is defined here as a collection of ideas that are at the same level of abstraction. Each idea group corresponds to a level of abstraction. In Figure 6, ideas a1, a2, and a3 form an idea group called A, ideas b1, b2, and b3 form an idea group called B, and so forth. For instance, suction, blowing, pushing, and pulling can together form an idea group at the phenomenon level and hose, collection chamber, and vacuum pump can together form an idea group at the part level.

The *size of an idea group* (s( )) is taken here as the number of ideas in that idea group. In Figure 6, s(A) = 3, s(B) = 3, and s(C) = 4.

An *idea space* (IS) is a collection of idea groups at all abstraction levels. Each group in an IS forms a level of abstraction and consists of a collection of distinct ideas, all at the same abstraction level. Thus, an IS consists of a collection of distinct ideas at different levels of abstraction. In Figure 6, IS consists of idea groups A, B, and C, which are characterized by their respective levels of abstraction. Groups A, B, and C individually consist of a collection of ideas {a1, a2, a3}, {b1, b2, b3}, and {c1, c2, c3, c4}, respectively, together constituting the IS. For example, an IS can consist of idea groups at action level {remove dust, transport dust, dispose dust}, state change level {change in state from dust to no dust, change in position of dust, change in color of dust}, input level {force, pressure, acceleration}, phenomenon level {suction, blowing, pushing, pulling}, effect level {Newton's second law of motion, Bernoulli's effect}, organ level {degrees of freedom of motion, value of mass, acceleration in direction of force}, and part level {hose, collection chamber, vacuum pump}.

A *concept space* for a function is a collection of alternative concepts that satisfy the function.

A *new concept space* (NCS) is a set of all concepts produced in a given design process that satisfy the same function. In Figure 6, C1, C2, C3, and C4 satisfy *F* and constitute an NCS.

An *existing concept space* (ECS) is defined as a collection of all concepts for a given function that existed even before the first solution in the NCS was designed. In Figure 6, E1, E2, and E3 constitute the ECS for *F*.

The *variety of a concept* (V) in a concept space is defined as a measure of the difference between the concept and all the other concept(s) produced previously in that concept space. For instance, in Figure 6, the variety of C4 (shown with dotted lines and double-headed solid arrow heads) is represented by comparing it with the previously produced concepts: C1, C2, and C3.

The *variety of a new concept space* [V(NCS)] is defined as the average of the values of variety of all the concepts in that concept space.

The *novelty of a concept* (N) in a concept space is defined as a measure of the difference between the concept and concepts in the ECS that satisfy the same function and concepts(s) previously produced in that concept space. For instance, in Figure 6, the novelty of C4 is depicted by comparing it against existing concepts E1, E2, and E3 (shown with dotted lines and double oval heads) and previously produced concepts C1, C2, and C3 (shown with dotted lines and double-headed solid arrow heads).

The *novelty of a new concept space* [N(NCS)] is defined as the average of the values of novelty of all the concepts in that concept space.

The *variety of an idea space* [V(IS)] is defined as a measure of the difference all the ideas from one another in that idea space. In Figure 6, the variety of IS (shown with dashed lines and double open-arrow heads) is depicted by comparing all of the ideas.

3.2. Research hypothesis

The hypothesis that we propose in this section relates novelty of a concept to the SAPPhIRE constructs that are used in designing that concept. The hypothesis therefore gives answers to the research questions raised in Section 2.8. A concept is made of ideas that are at different levels of abstraction (shown by single-headed solid arrow lines in Fig. 6). For the same function, the concepts in a NCS are different in terms of the ideas and the combination of ideas used. This contributes to the variety of the concept and the concept space. Because each concept is made of a number of ideas and the variety of each concept is brought out by ideas and combination of ideas used, variety of a NCS depends on the size of the idea groups and the V(IS). V(IS) also depends on the size of idea groups. Novelty of a concept depends on the difference of that concept from concepts in the ECS that satisfy the same function and concepts synthesized before in that new concept space. Novelty of a concept is also brought out by the ideas and the combination of ideas used. Therefore, it can be argued that N(NCS) depends on the V(NCS) and concepts in the ECS. However, concepts in the ECS cannot be controlled. Therefore, N(NCS) depends on the V(NCS), which is a function of the S(IS) and V(IS).

The hypothesis for this study can be thus stated as the following: *an increase in the number and variety of ideas explored while designing should enhance the variety of concepts produced, leading to an increase in the novelty of the new concept space.*

Each construct of SAPPhIRE constitutes an idea group. The use of the SAPPhIRE model can create multiple alternative ideas at each level of abstraction. This has been demonstrated with examples taken from multiple domains in Srinivasan and Chakrabarti (2009). This contributes to variety in an idea space. This can potentially increase the chances of producing a number of distinct alternative concepts that are different from each other at multiple levels of abstraction. This contributes to variety in a concept space. The variety in a concept space can potentially increase the chances of producing novel concepts. This contributes to novelty in a concept space. Thus, the hypothesis is a means to answer the research questions raised in Section 2.8. However, the hypothesis needs to be evaluated empirically to check its validity. A set of existing observational studies of designing sessions are used for the purpose of evaluating the hypothesis.

4. RESEARCH METHODOLOGY

To achieve the research objectives in the article through the hypothesis, eight existing observational studies of designing sessions are used. Each designing session involves an individual designer, experienced (E1–E4) or novice (N1–N4). The experienced designers have at least two and the novel designers have less than 2 years of designing experience. In each session, a designer solved a conceptual design problem (P1 or P2), under laboratory settings, by following a think-aloud protocol. Refer to Table 2 and Table 3, for the design problems and the problem-designer combinations used. In each case, the designer was instructed to generate as many conceptual-level solutions as possible, select one among them, and detail it; no time constraint was enforced. The designers were not told anything about the SAPPhIRE model, before and while designing. The designing sessions were videotaped, and the utterances and other actions of designers were transcribed. The designing sessions that were already available in a transcript form are used to answer the research questions.

The following steps are adopted:

1. *Identification of concepts*: All concepts generated in each designing session are identified from the transcripts. In each designing session, designers were instructed to generate as many concepts as possible, select one among them, and detail it. In this study, we consider all of the concepts that were generated.
2. *Identification of ideas*: All of the ideas in the transcript are identified by coding the transcript. The ideas are coded as one of the constructs of SAPPhIRE. Table 4 provides the symbols used in coding. Not all of the ideas are used in the generated concepts, but in this study we take all ideas into account, irrespective of whether they are used in the concepts. An intercoder reliability test is conducted to assess the subjectivity and ascertain whether the subjectivity is within the acceptable limits (refer to Appendix A).
3. *Estimation of the variety of concept*: The objectives in this article require an assessment of the variety of each concept, and accordingly an assignment of a proportionate number. To assess the variety, a number rating between 1 and 7 (both inclusive) is used corresponding to the seven SAPPhIRE constructs. Here, 7, 6, 5, 4, 3, 2, and 1 correspond to action, state change, input, phenomenon, effect, organ, and part levels, respectively. A value of 0 is assigned if there is no variety (refer to Appendix B). The following rules of thumb are followed while assigning a variety score to a concept:

   a. Irrespective of its nature, the first concept is always given a variety score of 0. This is so because variety requires at least two concepts.
   b. The second concept is compared with the first concept to ascertain the ideas that differentiate it from the first. The idea at the highest abstraction level is identified and a variety score is assigned based on the abstraction level of this idea. In this assessment, the idea at the highest abstraction level is considered, because a difference at a higher abstraction level would cause differences at all the subsequent lower levels of abstraction.
   c. The third concept is compared with the first and the second concept in the NCS to ascertain the ideas, that differentiates it from the first and second. The same procedure is repeated to assess its variety.
   d. In general, the nth concept is compared with all ($n - 1$) concepts ($n > 1$) generated previously in that concept space to ascertain the ideas that differentiate the nth concept from the others in that concept

Table 2. *Problems solved by designers*

Problem	Description
P1	Design a machine to make holes of any dimension in three dimensions, subject to the following constraints: (a) machine should be able to change the direction while the hole is being made; (b) machine should be able to make holes of different sizes; (c) machine should be able to make holes in metal, plastic, or wood; and (d) machine should be simple, small, and portable
P2	Design of a device to clean utensils subject to the following constraints: (a) device may be handheld or powered; (b) device is meant for urban middle class family with a maximum of 10 members; (c) should be able to clean all kinds of utensils like tumbler, dining plate, pressure cookers, and mixer grinder using this device; and (d) should be able to clean utensils made of all general kinds of materials like stainless steel, porcelain, glass, plastic, and aluminum

Table 3. *Observational studies showing the problem and designer combinations used*

	Designers							
	Experienced				Novice			
	E1	E2	E3	E4	N1	N2	N3	N4
Problem	P1	P2	P2	P1	P1	P1	P2	P2

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

Table 4. *Symbols used for coding*

	Construct						
	Action	State Change	Input	Phenomenon	Effect	Organ	Part
Symbol	a	s	i	ph	e	r	p

space. An idea corresponding to the highest level of abstraction is identified from among the ascertained ideas, and a variety score is assigned based on the abstraction level of this idea. This procedure is repeated until all the concepts are assigned a variety score.

4. *Computation of the variety of a new concept space*: The variety of an NCS of a designer is computed by averaging the variety score of all the concepts in that concept space. A simple average is calculated by adding the variety scores of all the concepts in the new concept space, and dividing the sum by the number of concepts in the new concept space.
5. *Estimation of an ECS*: An ECS is formed by collecting information from internet of products (see Appendix C) that perform similar or related functions as that given in the problem. It is quite difficult to form an exhaustive collection of existing concepts because there may be thousands of solutions that perform the same or similar function. In addition to the above information from the Internet, the knowledge of an experienced designer is also used.
6. *Estimation of the novelty of a concept*: The objectives in this article require an assessment of novelty, and accordingly an assignment of a proportionate number. To assess novelty of a concept, a number rating between 1 and 7 (both inclusive) is used corresponding to the seven SAPPhIRE constructs. Here, 7, 6, 5, 4, 3, 2, and 1 correspond to action, state change, input, phenomenon, effect, organ, and part levels, respectively. A value of 0 is assigned if there is no novelty (see Appendix B). The following rules of thumb are followed in assigning a novelty score to a concept:

   a. The first concept is compared with all the concepts in the ECS to ascertain the ideas that differentiate this concept. In addition, an experienced designer is consulted to cross-verify the ascertained ideas. The idea at the highest abstraction level from among the ascertained ideas is identified and a score is assigned, based on the abstraction level of this idea.
   b. The second concept is compared with the first concept and all the concepts in the ECS to ascertain the ideas that differentiates this concept. The experienced designer is consulted to cross-verify the ascertained ideas. The idea at the highest abstraction level is identified and a score is assigned based on the abstraction level of this idea.
   c. In general, the nth concept is compared with all the $(n - 1)$ concepts previously produced in the NCS as well as all the concepts in the ECS, to ascertain the ideas that differentiate the nth concept. This is cross-verified with the experienced designer. The idea at the highest abstraction level is identified among the ascertained idea, and a novelty score is assigned based on the abstraction level of this idea. The procedure is repeated, until all the concepts in the concept space have been assigned a novelty score.

   The procedure for estimating the variety and the novelty scores are the same, except that in novelty, in addition to comparing a concept with the previously generated concept(s), concepts in the ECS are also considered for the comparison.

7. *Computation of the novelty of a new concept space*: The novelty of the NCS of the designer is computed by averaging the novelty score of all the concepts in that concept space. A simple average is calculated by adding the novelty scores of all concepts in the new concept space and dividing the sum by the number of concepts in the new concept space.
8. *Computation of the size of an idea space*: The number of ideas at each level of abstraction is found by counting them individually. As mentioned earlier, not all ideas are used in the concepts by the designers. However, all the ideas produced in each designing session are taken into account.
9. *Computation of the variety of an idea space*: The variety of an idea space should take into account the following: variety is always proportional to number of ideas at any level of abstraction, that is, there is no variety if the number of ideas at an abstraction level is less than 2; and ideas at higher levels of abstraction should account for more variety than the ones at lower levels of abstraction. The following formula is developed based on these propositions:

$$V(IS) = \sum_{j=a}^{p} w_j(n_j - 1), \tag{4}$$

where n_a, n_s, n_i, n_{ph}, n_e, n_r, and n_p are the number of ideas at the action, state change, input, phenomenon,

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

effect, organ, and part levels, respectively, and w_a, w_s, w_i, w_{ph}, w_e, w_r, and w_p are the weightages at the action, state change, input, phenomenon, effect, organ, and part levels, respectively.

The weightages are arbitrarily assigned the following values: $w_a = 7$, $w_s = 6$, $w_i = 5$, $w_{ph} = 4$, $w_e = 3$, $w_r = 2$, and $w_p = 1$. In this formula, if the number of ideas at a particular abstraction level is 1 (i.e., $n_j = 1$), then this abstraction level does not contribute to variety because $(n_j - 1)$ becomes zero. If a particular abstraction level is neglected or not considered, that is, $n_j = 0$, then this abstraction level decreases the overall variety because $(n_j - 1)$ becomes negative. This serves as a form of penalty for skipping an abstraction level or jumping across an abstraction level.

10. *Computation of correlation values*: To determine if any relationship exists and the degree of relationships between variety or novelty and the different abstraction levels, the following variables are correlated and the degree of correlation is calculated:

    a. Variety of new concept space: size of idea group at various abstraction levels
    b. Novelty of new concept space: size of idea group at various abstraction levels

To test the hypothesis (proposed in Section 3.2), the following variables are correlated:

a. Variety of new concept space: variety of idea space
b. Novelty of new concept space: variety of new concept space

Pearson's correlation (Kreyszig, 2001) is used to compute the correlation values and it is computed through the in-built function in Microsoft Excel. All the correlation values are checked for their level of significance (Microbiologybytes, 2004) for the given number of observational studies analyzed. Level of significance for a variable is a measure of probability that a value for the given variable is accurate.

5. RESULTS

The results in this section are illustrated with examples taken from a designing session involving designer E2 who solved problem P2. E2 generated eight concepts. Some of the concepts are explained as follows (Table 5). In the first concept, the designer uses *pressurized cleaning* for cleaning the utensils. The *pressure* is sourced from an *external compressor*. In the second concept, the designer uses *abrasive cleaning* for cleaning the utensils. The designer feels that this concept is good for the harder materials but for the softer materials like stainless steel and aluminium, it may spoil the surface finish of the material. In the third concept, the designer uses a means of *dissolving in a solvent* to clean the utensils. This concept requires a *container* with a huge amount of *solvent*. The utensils are *dipped* in the solvent, *left* in the container for 5–10 min, taken out of the *container*, and *washed* under *flowing water*. The designer is not sure of the kind of solvent to be used but feels something like a detergent cannot be used for cleaning stubborn stains. According to the designer, the container cannot be hand held, but this concept requires less technical gadgetry.

The ideas in the first concept are pressurized cleaning (phenomenon), pressure (input) and compressor (part). The ideas in the second concept are abrasive cleaning (phenomenon). The ideas in the third concept are dissolving (phenomenon), container (part), solvent (part), and so forth. Even though it is expected that ideas from each abstraction level together make a concept, as observed from the above examples not all abstraction levels are used in constructing every concept. Table 6 shows an instance of each SAPPhIRE construct taken from the transcriptions. Refer to Appendix D for results of the intercoder reliability test. Table 7 shows the number of ideas produced by each designer in the designing sessions, at different levels of abstraction. In most cases, the number of ideas especially at the state change, phenomenon, effect, and organ levels are very low compared to those at the action and part levels. These results comply well with the findings reported in Srinivasan and Chakrabarti (2008), which used a different set of observational studies and yet arrived at similar results. Both of these sets of observational studies were undertaken before the SAPPhIRE model was developed.

Table 5. *Designer E2 comments while developing the concepts*

Concept	Utterances
C1	Pressure cleaning . . .
	For pressure what do we do?
	Normal household water pressure—can it be used?
	I don't think so; it needs to be pressurized.
	So we need a compressor, which is a big affair, hmm.
	So big constraint here is compressor.
C2	Abrasive cleaning . . . ,
	Abrasive is good for harder material but soft materials like aluminum or even stainless steel, it will spoil the finish of, may spoil the finish, . . .
C3	Dissolving in a solvent. . . . dissolving in a solvent . . . this does not spoil . . . this does not require any technical gadgetry but amount of solvent . . .
	Amount of solvent and also a container which sounds ok . . . could be one of the feasible things.
	A container [sketches it on the paper], wherein, this is not a handheld device [points at the sketch] . . . but it's ok . . . it can be handheld easily.
	Some . . . some vessel wherein you have some solvent, dip all the utensils, leave it for some time, maybe 5–10 minutes, take it out and put it under the flowing water and it [pointing at the vessel in the sketch] is cleaned.
	Could be . . . could be a nice solution but what will be this solvent?
	I . . . I don't know.
	Solvent . . . something like detergent maybe but detergents are not that effective in cleaning those stubborn stains.
	For the time being let me assume that I don't know the solvent . . . that could be some solvent available which can clean.

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

Table 6. *Instance of SAPPhIRE constructs from transcriptions*

Construct	Instance From the Transcription
a	*Third will be cleaning with the detergent.* [This step is used as a third step in a series of steps to clean the dirty utensil. "Cleaning" is taken as the action.]
s	*The material of this which was here is now removed.* [In this instance, the designer describes a consequence of a means to remove the material; the consequence is the state change: material → no material.]
i	*The jet of water is impacted on the surface and the force of water removes the material.* [In this instance, the designer uses a jet of water to remove material from a block; the designer further states that the force of water is used to remove the material. "Force" is taken as the input.]
ph	*Thirdly it can be like that . . . tilting it (dirty dish) and jerking it (dirty dish).* [The designer uses tilting of the dirty dish, followed by jerking the dirty dish to remove the leftovers from the dish. "Tilting" and "jerking" are taken as phenomena.]
e	*I use the principle of magnetism to control the direction of flow of chemical, which removes the material.* [In this instance, the designer uses the principle of magnetism to guide the liquid through the block in the direction required for the hole to be made; the liquid is used to remove the material of the block. "Magnetic principle" is taken as the effect.]
r	*So, you may have some sort of flexible hose.* [The designer decides to use a flexible hose for transporting the liquid inside the block in which a hole is desired to be made; the liquid is used to remove the material of the block. The "flexible" property of the hose is taken as the organ.]
p	*Use local heating/burning to remove material using a heat gun.* [The designer generates an idea of using a heat gun to locally heat or burn the material of the block, thereby making a hole. "Heat gun" is taken as the part.]

The first concept irrespective of its nature is assigned a variety score of zero, for reasons explained earlier in Section 4 under point 3a. The first concept, which uses pressurized cleaning, already exists in the existing concepts, and therefore, there is no difference between the first concept and the existing concepts. Hence, a novelty score of 0 is assigned. The highest difference between the second concept and the first concept is at the phenomenon level (first concept uses pressurized cleaning and second concept uses abrasive cleaning), and therefore, a variety score of 4 is assigned. The highest difference of the second concept from the first concept and the ECS is at the phenomenon level (second concept uses abrasive cleaning, which is not used in the first and existing concepts), and therefore, a novelty score of 4 is assigned. The highest difference of the third concept from the first and second concepts is at the action level (third concept uses washing by means of flowing water after dissolving using solvent; washing is a more general form of pressurized cleaning, abrasive cleaning, and flowing water-type cleaning used in the first, second, and third concepts, respectively), and therefore a variety score of 7 is assigned. The highest difference of the third concept from the first, second, and existing concepts is at the phenomenon level (dissolving is not used in the first, second, and existing concepts) and therefore, a novelty score of 4 is assigned. The procedure is repeated for estimating the variety and novelty scores for the other concepts in the NCS of the designer.

An average of the scores of variety and novelty of all the concepts in the NCS respectively constitute the variety and novelty of that concept space. Table 8 shows the value of variety and novelty of NCS for each session. Theoretically, the value of novelty should be less than that for variety, because novelty accounts for concepts in the existing and new concept space. A similar trend is observed for all the eight cases.

Before variety/novelty values are correlated with the size of idea groups at different abstraction levels, the idea groups are combined. Actions, state changes, and inputs; phenomena and effects; and organs and parts, are taken as the three combined groups. This is done for the following reasons: (a) designers in the analyzed sessions were allowed to work in a natural way and not instructed about SAPPhIRE before or while designing and there were thus less descriptions at the s, i, e, and r levels; and (b) a-, s-, and i-level descriptions together constitute the functional-level description; ph and e levels together constitute the behavioral-level description; and r and p levels together constitute the structural-level description, enabling comparisons between variety/novelty with the size of functional-, behavioral-, and structural-level descriptions. This allows the comparisons to be more generic

Table 7. *Number of ideas at different abstraction levels*

SAPPhIRE Construct	Experienced				Novice			
	E1 (P1)	E2 (P2)	E3 (P2)	E4 (P1)	N1 (P1)	N2 (P1)	N3 (P2)	N4 (P2)
a	9	8	7	6	6	7	13	12
s	1	0	0	1	2	0	0	0
i	1	1	1	2	0	0	0	1
ph	32	7	12	5	9	3	11	7
e	1	0	0	0	0	0	2	0
r	19	2	1	4	1	2	5	1
p	40	20	16	14	25	9	9	18

Table 8. *Variety (V) and novelty (N) of new concept space (NCS)*

	Experienced				Novice			
	E1 (P1)	E2 (P2)	E3 (P2)	E4 (P1)	N1 (P1)	N2 (P1)	N3 (P2)	N4 (P2)
V(NCS)	4.44	3.88	3.75	3.00	2.42	3.14	4.54	3.69
N(NCS)	3.89	3.13	2.92	2.57	1.58	2.14	4.00	3.54

and the results to be valid in a more generic sense. Table 9 shows the size of the combined idea groups.

Table 10 shows the correlation between variety/novelty of NCS with the size of combined idea groups (a + s + i/ph + e/r + p). Note that the number in each cell of the table represents the correlation value between the row and column that connects the cell. The number inside the corresponding bracket represents the level of significance for the sample size ($n = 8$). For instance, 0.66 represents the correlation value between V(NCS) and s(a + s + i) and the significance value falls in the range of 0.90–0.95. The following observations can be drawn:

1. Correlation values between the variety of the NCS and the size of (actions, state changes and inputs), (phenomena and effects), and (organs and parts) are in descending order. This signifies that the variety of the NCS is proportional to the abstraction level of the constructs that are explored.
2. Correlation values between novelty of the NCS and the size of (actions, state changes and inputs), (phenomena and effects), and (organs and parts) are in descending order. This signifies that novelty of the NCS is also proportional to the abstraction level of the constructs that are explored.
3. Variety and novelty are computed using the procedure explained in Section 4, under points 3 and 6. The method used assigns higher variety/novelty scores when constructs at higher abstraction levels are used. The results observed in this section in points 1 and 2 should not be attributed to the above cause. This is so, because the size of idea groups takes into account ideas that are used in the concepts, as well as those that are not used.
4. The above observations are valid findings because any change in a higher abstraction level has a greater chance of producing a highly different concept, that is, higher chances of variety/novelty. This is a result of the fact that a change in a higher abstraction level can potentially cause changes in the all subsequent lower abstraction levels.

Table 9. *Size of combined idea groups for idea space (IS)*

	Experienced				Novice			
s(IS)	E1	E2	E3	E4	N1	N2	N3	N4
s(a + s + i)	11	9	8	9	8	7	13	13
s(ph + e)	33	7	12	5	9	3	13	7
s(r + p)	59	22	17	18	26	11	14	19

Table 10. *Correlation values between variety (V) or novelty (N) of new concept space (NCS) and size of idea groups*

	V(NCS)	N(NCS)
s(a + s + i)	0.66 (0.90–0.95)	0.82 (0.98–0.99)
s(ph + e)	0.60 (<0.90)	0.56 (<0.90)
s(r + p)	0.33 (<0.90)	0.33 (<0.90)

Table 11 shows the correlation values between the variety of the NCS and the variety of corresponding idea space or NCS. The number inside the corresponding bracket represents the level of significance for the sample size ($n = 8$). The values support the hypothesis, that is, an increase in the variety of ideas explored increases the variety of concepts, which in turn increases novelty of concepts.

Correlation values that fall in the ranges of 0.3–0.5 and 0.5–1.0 are taken to be medium and large, respectively, in Wikipedia (2009).

6. DISCUSSION

This section discusses the work in this article in the context of the existing literature. Shah et al. (2003) proposed four outcome-based measures for measuring a designer's ideation: variety, novelty, quantity, and quality. An idea and a solution as used by Shah et al. (2003) are similar respectively to a concept and an idea as used in our work. Shah et al. (2003) stated that if two ideas use different physical principles for satisfying the same function, then the ideas are very different from each other; in contrast, if two ideas are different from each other in only a secondary construction level of detail, then the ideas are only slightly different from each other. Shah et al. (2003) used weights p_k and s_k in Eqs. (1) and (3) for calculation of novelty and variety; a higher stage has a higher score. This resembles our work where a higher score of variety and novelty is assigned if the highest difference between two or more con-

Table 11. *Correlation value to validate hypothesis for variety (V) or novelty (N) of new concept space (NCS) or idea space (IS)*

	V(NCS)
V(IS)	0.65 (0.90–0.95)
N(NCS)	0.95 (>0.99)

cepts is at a higher abstraction level. Variety and quantity as used by Shah et al. (2003), if combined, resembles the variety of the idea space used in our work. There are differences between the work of Shah et al. (2003) and ours. Shah et al. (2003) stated that while computing the variety, if the number of branches (i.e., the number of solutions at an abstraction level) is 1, then the variety score (at this level) is taken to be 0 while using it in the formula for computing variety, although the formula itself does not account for this. An important case where the number of outcomes at an abstraction level can be 0 (designers often skip abstraction levels, and thus the number of outcomes at an abstraction level can be 0) is unlikely to be covered by Eq. (3). However, the equation for computing variety of idea-space in this article resolves the above two issues by incorporating a factor of $(n_j - 1)$ [refer to Eq. (4)]. Thus, if the number of ideas at an abstraction level is 1 or 0, $(n_j - 1)$ would result in zero or negative contribution to variety, respectively.

Several researchers (Osborn, 1979; Chakrabarti and Bligh, 1994; Candy, 1996; Cross, 1996; Shah et al., 2003) argued that generating several ideas would increase the chances of producing better ideas. Kurtoglu et al. (2009) argued that increasing the number of solution principles and components in their knowledge base improved the degree of variety and novelty, where the knowledge base has been used in a computational tool for generating conceptual design solutions for electromechanical problems. However, none of the above approaches verified this broad hypothesis that an increase in the number and variety of ideas influence the novelty of concepts. In our article, a similar approach is proposed, and also empirically verified. We argue that an increase in the number of solution principles is the result of exploring a larger variety (which in our measure includes both the number of and differences between ideas) of ideas. Our hypothesis relates the variety of ideas explored through the variety of concepts generated to the novelty of concepts, and is verified empirically using protocol studies of designing sessions.

7. SUMMARY AND FUTURE WORK

The section presents a summary of the article and some directions for future work.

1. The research in this article is an attempt at understanding the relationship between novelty of a concept and the product facets (outcomes) that are used in designing the concept. This relationship is constructed with a hypothesis: an increase in the variety of ideas explored while designing should enhance the variety of concepts, leading to an increase in the novelty of the concepts.
2. The hypothesis is verified empirically using existing observational studies of designing sessions. In the sessions, designers did not make explicit use of the SAPPhIRE constructs.
3. Results revealed that novelty and variety of the NCS are directly related to abstraction levels. This infers that there is a greater chance of designing a novel concept, if the higher abstraction levels are explored in more detail. This inference gains importance especially when designers do not use adequate numbers of laws and effects in designing; significant quantity of novelty may be lost. This issue calls for a need to support designers with knowledge of laws and effects in designing.
4. This research establishes a relationship between novelty of an NCS and the constructs of the SAPPhIRE model explored while designing concepts in that concept space. However, a more comprehensive evaluation of the model can be done by comparing concepts produced by designers without the model and with the model.

REFERENCES

Amabile, T.M. (1996). *Creativity in Context: Update to the Social Psychology of Creativity*. Westview, CO: Westview Press.

Andreasen, M.M. (1980). *Syntesemetoder pa systemgrundlag*. PhD Thesis. Lund Technical University, Sweden.

Berliner, C., & Brimson, J.A. (1988). *Cost Management for Today's Advanced Manufacturing: The CAM-I Conceptual Design*. Cambridge, MA: Harvard Business School Press.

Blessing, L., & Chakrabarti, A. (2009). *DRM, a Design Research Methodology*, pp. 120–121. London: Springer.

Blessing, L., Chakrabarti, A., & Wallace, K. (1995). A design research methodology. *Proc. Int. Conf. Engineering Design*, pp. 502–507.

Boden, M.A. (1999). Computer models of creativity. In *Handbook of Creativity* (Sternberg, R.J., Ed.), pp. 351–372. New York: Cambridge University Press.

Cambridge University Press. (2009). *Cambridge Dictionaries Online.* New York: Cambridge University Press. Accessed at http://www.dictionary.cambridge.org/

Candy, L. (1996). Understanding creativity: an empirical approach. *Proc. Creativity and Cognition*, pp. 45–54.

Chakrabarti, A. (2002). Editor's introduction. In *Engineering Design Synthesis: Understanding, Approaches and Tools* (Chakrabarti, A., Ed.), pp. vii–xvi. London: Springer–Verlag.

Chakrabarti, A. (2003). Towards a measure for assessing creative influences of a creativity technique. *Proc. Int. Conf. Engineering Design* [CD].

Chakrabarti, A., & Bligh, T. (1994). Functional synthesis of solution-concepts in mechanical conceptual design. Part I: knowledge representation. *Research in Engineering Design 6(3)*, 127–141.

Chakrabarti, A., & Khadilkar, P. (2003). A measure for assessing product novelty. *Proc. Int. Conf. Engineering Design* [CD].

Chakrabarti, A., Johnson, A., & Kiriyama, T. (1997). An approach to automated synthesis of solution principles for microsensor designs. *Proc. Int. Conf. Engineering Design*, pp. 125–128.

Chakrabarti, A., Sarkar, P., Leelavathamma, B., & Nataraju, B.S. (2005). A functional representation for aiding biomimetic and artificial inspiration of new ideas. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 19(2)*, 113–132.

Cross, N. (1996). Creativity in design: not leaping but bridging. *Proc. Creativity and Cognition*, pp. 27–35.

Eder, W.E. (1995). Introduction. *Proc. Int. Workshop on Engineering Design and Creativity*, pp. 16–18.

Franzosi, M. (2006). Novelty and non-obviousness—the relevant prior art. Accessed at www.law.washington.edu/casrip/Symposium/Number7/

French, M. (1999). *Conceptual Design for Engineers*, 3rd ed. Berlin: Springer–Verlag.

Hubka, V. (1976). *Theorie der Konstruktionsprozesse*. Berlin: Springer–Verlag.

Koyama, T., Taura, T., & Kawaguchi, T. (1996). Research on natural law database. *Proc. Joint Conf. Knowledge Based Software Engineering*, pp. 242–245.

Kreyszig, E. (2001). *Advanced Engineering Mathematics*, 8th ed. Singapore: Wiley.

Kurtoglu, T., Campbell, M., & Linsey, J. (2009). An experimental study on the effects of a computational design tool on concept generation. *Design Studies 30(6)*, 676–703.
Lopez-Mesa, B., & Vidal, R. (2006). Novelty metrics in engineering design experiments. *Proc. DESIGN*, pp. 557–564.
Microbiologybytes. (2004). Critical values of correlation coefficient. Accessed at http://www.microbiologybytes.com/maths/rtable.html
Molina, A., Al-Ashaab, A.H., Ellis, T.I.A., Young, R.I.M., & Bell, R. (1995). A review of computer-aided simultaneous engineering systems. *Research in Engineering Design 7(1)*, 38–63.
Murakoshi, S., & Taura, T. (1998). Research on the systematization of natural laws for design support. *Proc. 3rd IFIP Workshop on Knowledge Intensive CAD*, pp. 141–160.
Osborn, A. (1979). *Applied Imagination*. New York: Scribners.
Ottosson, S. (1995). Boosting creativity in technical development. *Proc. Workshop on Engineering Design and Creativity*, pp. 35–39.
Pahl, G., & Beitz, W. (1996). *Engineering Design: A Systematic Approach*. London: Springer–Verlag.
Sarkar, P. (2007). *Development of a support for effective concept exploration to enhance creativity of engineering designers.* PhD Thesis. Indian Institute of Science.
Sarkar, P., & Chakrabarti, A. (2007*a*). Understanding search in design. *Proc. Int. Conf. Engineering Design*, pp. 319–320.
Sarkar, P., & Chakrabarti, A. (2007*b*). Development of a method for assessing design creativity. *Proc. Int. Conf. Engineering Design*, pp. 349–350.
Sarkar, P., & Chakrabarti, A. (2008). Studying engineering design creativity. *Workshop on Studying Design Creativity*, France.
Saunders, R. (2002). *Curious design agents and artificial creativity—a synthetic approach to the study of creative behaviour.* PhD Thesis. University of Sydney, Faculty of Architecture.
Shah, J.J., Smith, S.M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. *Journal of Design Studies 24(2)*, 111–134.
Shalley, C.E. (1991). Effects of productivity goals, creativity goals, and personal discretion on individual creativity. *Journal of Applied Psychology 76(2)*, 179–185.
Srinivasan, V., & Chakrabarti, A. (2008). Design for novelty—a framework? *Proc DESIGN*, pp. 227–232.
Srinivasan, V., & Chakrabarti, A. (2009). SAPPhIRE—an approach to analysis and synthesis. *Proc. Int. Conf. Engineering Design* [CD].
Sternberg, R.J., & Lubart, T.I. (1999). The concept of creativity: prospects and paradigms. In *Handbook of Creativity* (Sternberg, R.J., Ed.), pp. 3–15. New York: Cambridge University Press.
Thring, M.W., & Laithwaithe, E.R. (1977). *How to Invent*. New York: Macmillan Company.
Umeda, Y., Ishii, M., Yoshioka, M., Shimomura, Y., & Tomiyama, T. (1996). Supporting conceptual design based on the function–behavior–state modeler. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 10(4)*, 275–288.
Unsworth, K.L. (2001). Unpacking creativity. *Academy of Management Review 26(2)*, 286–297.
Westwood, A.R.C., & Sekine, Y. (1988). Fostering creativity and innovation in an industrial R & D laboratory. *Research-Technology Management 31(4)*, 16–20.
Wikipedia. (2009). Correlation. Accessed at http://en.wikipedia.org/wiki/Correlation
Yoshioka, M., & Tomiyama, T. (1997). Pluggable metamodel mechanism: a framework of an integrated design object modelling environment. *Proc. Lancaster Int. Workshop on Engineering Design CACD '97*, pp. 57–70.
Zavbi, R., & Duhovnik, J. (2000). Conceptual design of technical systems using functions and physical laws. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 14(1)*, 69–83.
Zavbi, R., & Duhovnik, J. (2001). Analysis of conceptual design chains for the unknown input/known output pattern. *Proc. Int. Conf. Engineering Design*, pp. 53–60.
Zimmerman, L.W., & Hart, G.D. (1998). *Value Engineering—A Practical Approach for Owners, Designers and Contractors*. New York: Van Nostrand Reinhold.

V. Srinivasan is a PhD student at the Centre for Product Design and Manufacturing at the Indian Institute of Science, Bangalore, India. He holds a BEng degree in mechanical engineering from the University of Madras, Chennai, India. His research interests include design theory and methodology, design synthesis, design creativity, designing from first principles, and working of sensors and transducers.

Amaresh Chakrabarti is a Professor at the Centre for Product Design and Manufacturing at the Indian Institute of Science, Bangalore, India. He holds a PhD in engineering from the University of Cambridge. His research interests include functional synthesis, design creativity, design methodology, collaborative design, eco-design, engineering design, design synthesis, requirements management, knowledge management, computer-aided design, and design for variety.

APPENDIX A: INTERCODER RELIABILITY TEST

Blessing and Chakrabarti (2009) suggested the use of intercoder reliability test to assess the subjectivity of coding qualitative data, such as a transcript of a designing session. In our experience, the entire transcription has never been recoded by other coders for the purpose of testing subjectivity. Instead, some representative portions of transcription are given for recoding. The intercoder reliability test for our research is conducted in the following way:

1. For one of the design problems from this study, the problem, a portion of the transcription from a designing session of the problem, the different coding categories, and their definitions are given to three coders (Co1–3). The coders are asked to independently code the transcription by following these steps:
   a. Identify each outcome as solution (idea) or nonsolution for the given design problem.
   b. Categorize each outcome into one of the SAPPhIRE constructs (note that the nonsolutions need not be categorized).
2. The codes of these coders are individually compared with the actual codes (coding done by the authors of this article) to ascertain if there are any differences. The differences in codes are assessed in terms of
   a. identification of solution or nonsolution, and
   b. categorization of the solution into one of the SAPPhIRE constructs.
3. A comparison score for a coder is given based on the following rules:
   a. One point is assigned if there is a match in the identification of a solution or nonsolution.
   b. One point is assigned if there is a match in the categorization of the identified solution. (It has to be noted that a match in the categorization assumes that there is a also a match in the solution identification.)
   c. The comparison score for each coder is calculated by adding the total number of points, that is, the total number of matches. This score is referred to as the comparison score before intervention. A percentage comparison score is computed by dividing this comparison score by the maximum comparison score possible and multiplying by 100. The

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

percentage comparison score reflects the percentage agreement, which is a measure of intercoder reliability.

4. The first author of this article has a discussion with the three coders individually to exchange the rationales behind their respective codes. This is followed by a session of recoding of the same transcription section by the three coders.
5. The recoded codes are now compared with the actual codes to estimate the comparison score by following the same procedure explained in Step 3. This score is referred to as the comparison score after intervention. A percentage comparison score is computed by dividing this comparison score by the maximum comparison score possible and multiplying by 100.

APPENDIX B: METHOD TO ASSESS VARIETY/NOVELTY

The novelty assessment method proposed by Sarkar and Chakrabarti (2007*b*, 2008) explained earlier in Section 2.7, is modified to assess variety/novelty in this research. As explained earlier, the method proposed earlier uses the constructs of FBS and SAPPhIRE models. FBS constructs are used to assess novelty, and if the product is novel, then SAPPhIRE is used to assess the qualitative degree of novelty. This method does not use behavior and action of FBS and SAPPhIRE, respectively, and combines constructs—state change with input, phenomenon with effect, and organs with parts—of SAPPhIRE while assessing the degree of novelty.

The modified method uses only the constructs of SAPPhIRE and treats each of them separately in decreasing level of abstraction. The actions of the concept are identified and compared to check if they exist. If they do not exist then a score for variety or novelty is assigned. If they already exist, then state changes of the concept are identified and compared. If these do not exist, then a score for variety or novelty is assigned. If the state changes also exist, then the inputs of the concept are identified and compared. If these do not exist, then a score for variety/novelty is assigned. If these also exist, then phenomena of the concept are identified and compared. This method starts by comparing the constructs at higher levels of abstraction and proceeds in decreasing levels of abstraction: action, state change, input, phenomenon, effect, organ, and part. If parts of the product or concept also exist, then the product/concept does not have variety or novelty. This method continues until a score for variety/novelty is assigned.

For estimating variety, all concepts produced previously in the concept space are considered. The first concept is assigned a variety score of zero. The second concept is compared against the first concept. The differences are identified and based on a difference at the highest level of abstraction, a variety score is assigned. The third concept is compared against the first and second concepts, and the same procedure is repeated. The steps are repeated until all the concepts in the concept space are each assigned a variety score.

For estimating novelty, all concepts produced previously in the concept space and concepts in the ECS that satisfy the same functionality, are considered. The first concept is compared against all the concepts in the ECS. The differences are identified, and based on a difference at the highest level of abstraction, a novelty score is assigned. The second concept is compared against the first concept and all the concepts in the ECS. The differences are identified, and based on a difference at the highest level of abstraction, a novelty score is assigned. The third concept is compared against the first concept, the second concept, and the ECS, and the same procedure is repeated. The steps are repeated until all the concepts in the concept space are each assigned a novelty score.

The abstraction levels are assigned a score as shown in Table B.1; higher scores are assigned to higher abstraction levels and the score reduces as the abstraction level reduces. The reasoning for the score can be explained as follows: let us assume an equal weightage (say, a score of 1) for all of the abstraction levels. If a concept is different from earlier concepts (produced earlier in the same concept space and existing concept space) at a particular abstraction level (say, action level), this difference also creates differences at all of the lower abstraction levels (state change, input, phenomenon, effect, organs, and parts). Thus, when the difference of a concept from previous concepts is ascertained at the action level, its difference at the action level as well as all subsequent abstraction levels have to be accounted for (action + state change + phenomenon + effect + organs + parts, i.e., $1 + 1 + 1 + 1 + 1 + 1 + 1 = 7$). The same logic applies for scores for the other abstraction levels.

The above explanation is supported, for example, by the concepts of designers E4 and E1. In the first concept of designer E4, after the *work-piece is located* (A), *material removal* (A) takes place by converting it into an amorphous powder (S), which can be facilitated by charring (ph). This can be accomplished by supplying heat (I) that can be achieved by a laser (I) from laser-emitting heads (P), and heat is produced when two lasers meet at a point (R). In the second concept, after the *work-piece is located* (A), *material removal* (A) takes place by *drilling* (Ph) it. This form of drilling requires a *flexible* (R) *arm* (P) with a *drilling head* (P) *attached to its end* (P) to make holes of different sizes. In the estimation of variety, the first concept is taken as the reference. The second concept of E4 is compared with the first concept. Both of these concepts have the same action-level description. The second concept does not have any explicit state change- or input-level description. The second concept is different from the first in terms of phenomenon-, organ-, and part-level descriptions. The third concept is compared to the first and second concepts and so forth. A procedure similar to that explained in Section 4 under point 3 is used in assessing the variety of each concept. Table B.2 shows the differences. Concept C5 has the same input-change level description as C1 but uses a different phenomenon. The table also shows that designers do not explore all of the abstraction levels explicitly. Similar trends are also seen for designer E1 (Table B.3). An analysis of the concepts of designers shows that a difference at the highest level of abstraction ensures differences at all subsequent levels of abstraction. This empirical observation is formalized in the scoring scheme in a uniform manner across all levels, regardless of whether a designer explicitly describes every level of abstraction of each concept.

Table B.1. *Score for each abstraction level of SAPPhIRE*

Construct	Score
Action	7
State change	6
Input	5
Phenomenon	4
Effect	3
Organ	2
Part	1

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press

Table B.2. *Differences between concepts of designer E4*

	C1	C2	C3	C4	C5
A		Sa	Sa	Sa	Sa
S		—	—	—	—
I		—	—	—	Sa
Ph	Datum	Di	Di	Di	Di
E		—	—	—	—
R		Di	Di	—	—
P		Di	—	Di	Di

Note: Sa, same; Di, different; —, no description available.

APPENDIX C: EXISTING PRODUCTS

For problem P1:

http://computer.howstuffworks.com/stereolith.htm
http://en.wikipedia.org/wiki/Bulk_micromachining
http://en.wikipedia.org/wiki/Surface_micromachining
http://en.wikipedia.org/wiki/Etching
http://en.wikipedia.org/wiki/Dry_etching

For problem P2:

http://en.wikipedia.org/wiki/Dishwasher
http://home.howstuffworks.com/dishwasher.htm
http://www.apwagner.com/appliancerepaircenter/Dishwashers.htm#HowdoDishwasherswork

APPENDIX D: RESULTS OF INTERCODER RELIABILITY TEST

A portion of the transcription from a designing session of Problem P2 solved by Designer E3 consisting of 35 outcomes is given to three coders. Not all the outcomes in the transcript are solutions. Because there are 35 outcomes in the given transcription and each outcome has two subsections—solution/nonsolution identification and categorization of solution—the maximum comparison score is seventy (i.e., $35 \times 2 = 70$).

Table D.1 shows the comparison score and the percentage comparison score of each researcher under the column "before intervention." The lack of matches was primarily because of the following:

1. Coders are confused between solutions and requirements (a form of nonsolution). In one instance, when designer E3 uttered: "Dish washing process basically will be . . . I mean . . .(writes on the sheet of paper)," a coder took "washing of the dishes" as a solution instead of a requirement.
2. In some cases, the coders did not identify solutions and thus skipped their categorization. For example, when the designer uttered "First will be like . . . what you do . . . you put the dirty dishes in the sink," a couple of coders identified "put" as solution and categorized it as an action-level solution but missed "sink," and thus failed to categorize it.
3. In some cases, the coders are confused between SAPPhIRE constructs, particularly action and phenomenon. In one of the utterances of the designer: "Fourth will be rinsing it (dirty plate) and putting it (rinsed plate) back," coders categorized "rinsing" as an action rather than a phenomenon.

Table B.3. *Differences between concepts of designer E1*

	C1	C2	C3	C4	C5	C6	C7	C8	C9
A		Sa	Di	Sa	Sa	Sa	Sa	Di	Sa
S		Di	—	—	—	—	—	—	—
I		—	—	—	—	—	—	Di	—
Ph	Datum	—	—	Di	Di	Di	—	—	Di
E		—	—	—	—	—	Di	—	—
R		Di	—	Di	Di	—	—	Di	—
P		Di	Di	Di	Di	Di	Di	Di	Di

Note: Sa, same; Di, different; —, no description available.

Table D.1. *Comparison scores before and after intervention*

	Before Intervention			After Intervention		
	Co1	Co2	Co3	Co1	Co2	Co3
Comparison score (max. 70)	53	58	61	69	70	67
Comparison score percentage (max. 100)	75.71	82.86	87.14	98.57	100	95.71

During the discussion session with the first author of this article, the following points were highlighted:

1. The coders are instructed to read the objective of the design exercise again. This was meant to help them differentiate between nonsolutions and solutions.
2. The definition of solution and constructs of SAPPhIRE are explained with more examples from another study that was carried out in Srinivasan and Chakrabarti (2009). This was meant to help them gain a better understanding of the definition of solution and constructs of SAPPhIRE, so as to help them in better categorization.

The discussion session is followed by a session of recoding. The recoded transcription section is now compared with the actual codes to estimate the comparison score. Table D.1 shows the comparison score and the percentage comparison score under the column "after intervention." It is observed that there is a marked improvement in the comparison score "before intervention" and "after intervention." This shows that the discussion session was useful and helped to bridge the gap in subjectivity in coding of the transcription.

Blessing and Chakrabarti (2009) described a similar measure for intercoder reliability (defined as the ratio of number of agreements to the sum of number of agreements and number of disagreements) to ascertain the agreement between the coded transcriptions. They considered an agreement score of greater than 70% as acceptable. Because all the percentage comparison score in all the cases for the three coders here is more than 70%, no changes are made in the actual codes of the transcription.

https://doi.org/10.1017/S089006041000003X Published online by Cambridge University Press