Design Science

Assessment of the creative potential of design problems via novelty and usefulness

Sanjay Singh and Amaresh Chakrabarti

Department of Design and Manufacturing, Indian Institute of Science, Bangalore, India

Abstract

Current design theories and models predominantly focus on evaluating innovation through design solutions, using measures of novelty and usefulness as indicators of creativity. In contrast, the assessment of creative potential of design problems has attracted far less attention. To systematically explore the creative potential of design problems, a comprehensive literature review is conducted, revealing significant gaps where existing methods have yet to be applied. To address these gaps, first, an extensive database of design problems has been constructed using data collected from design patents, surveys, and questionnaires. Three distinct quantitative methods have been developed: the first for assessing novelty using SAPPhIRE model of causality, the second for assessing usefulness using usefulness indicators, and the third for assessing creative potential. The novelty method quantifies the minimum distance between a current problem and the old problems in the database, using textual similarity at different levels of SAPPhIRE abstraction. Expert evaluation of the novelty method indicates substantial agreement with experts' intuitive notion, in addition to higher effectiveness compared to existing methods. The first two methods have then been integrated into the third method for assessing the overall creative potential of a design problem. Statistical analyses confirmed the correlation of both novelty and usefulness with creative potential, supporting findings in the literature. To demonstrate the methods, detailed case studies have been presented, illustrating the application of the methods. This systematic approach provides a robust framework for objective assessment of creativity in design problems, facilitating better prioritization and decision-making in engineering design contexts.

Keywords: problem SAPPhIRE, creative potential, problem novelty, usefulness, multi-instance SAPPhIRE

Received 16 May 2025
Revised 05 November 2025
Accepted 06 November 2025

Corresponding author
Sanjay Singh
sanjaysingh@iisc.ac.in

Des. Sci., vol. 11, e51
journals.cambridge.org/dsj
DOI: 10.1017/dsj.2025.10043

the Design Society a worldwide community

CAMBRIDGE UNIVERSITY PRESS

1. Introduction

Engineering design is broadly divided into four stages: task clarification, conceptual design, embodiment design and detailed design (Pahl & Beitz 1996). In today's world, engineering designers encounter a wide range of design problems during the task clarification phase, subsequently generating multiple solutions throughout the rest of the three stages (Roozenburg & Eekels 1995). These may involve creating entirely new products or modifying existing ones (Siddharth & Sarkar 2018). In both scenarios, various solution alternatives are produced, assessed, refined and ultimately chosen (Srinivasan & Chakrabarti 2010a). Only those products that demonstrate creativity are likely to remain competitive in the marketplace (Hart &

Christensen 2002). Therefore, evaluating "creativity" is crucial when selecting solutions, as it helps identify those aspects of a product that lack creativity. Both engineering design and creative process are planned and occur together in practice (Basadur *et al.* 2000). Engineering design is a complex process of both problem finding and problem-solving. Various design studies have shown that participants invest little time in problem finding or understanding, which is the very first step of the design process, compared to that in problem-solving (Nidamarthi, Chakrabarti & Bligh 1997). Also, during the selection of design problems, the participants lean toward the selection of "problems that they believed in intuitively or found obvious, rather than those having creative potential."

The fundamental assumption in our approach is the following: "Identification of problems with greater creative potential as well as developing more creative solutions for them should lead to more creative solutions than when developing creative solutions for problems identified without attention to their creative potential." To justify this, we argue that typically, when people focus on problem-solving, they tend to take the problem for granted. One looks at the solution variants to see how those problems are solved. One does not ask the question, "Is this the most creative problem to solve?" So, potentially one layer of opportunity to be creative is lost, where one could ask that there be a number of possible variants of a design problem, of which some are more creative than others, and therefore, to maximize creativity, one should identify and focus on the one that is potentially the most creative. The focus then is on finding the different possible solutions to the problem and selecting the one that offers the most creative solution to the problem with the most creative potential (Runco & Chand 1994). This compounds creativity, as opposed to when identifying the problems without paying attention to their creative potential (Basadur *et al.* 2000).

An exploration of the creativity literature suggested the need for empirical studies on existing creativity assessment methods for the proposal of new or modified creativity assessment method for design problems. So, this manuscript proposes a method for assessing the creative potential of design problems, via assessment of their novelty and usefulness, during the task clarification stage in the engineering design process. It intends to help designers prioritize and focus on selecting problems with high novelty and (potential) usefulness, thereby enhancing their creative efforts beforehand and saving time and effort during the design process. Case studies are used to illustrate the applicability, which is then backed up by statistical analysis after comparison with existing methods and, finally, by experts' intuitive understanding of a sample of design problems.

Assessment of the novelty of design problems was previously attempted in our earlier work (Singh & Chakrabarti 2025) titled "Supporting Assessment of Novelty of Design Problems Using Concept of Problem SAPPhIRE," where a methodology for the assessment of the novelty of design problems was proposed and demonstrated with some examples, even though it was limited to fewer (only two sets) and simpler (with single-instance problem SAPPhIREs (State Changes, Actions, Parts, Phenomena, Inputs, Organs, Effects)) problems. No validation of the proposed novelty assessment method was carried out. It also lacked a method for assessment of (potential) usefulness of (solving a) design problem, as well as assessment of the

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

overall creative potential of a problem. The novelty of the current contribution in comparison with the earlier work is as follows:

(1) The novelty assessment method from previous work has been expanded to address more complex problems (represented using multi-instance SAPPhIRE models), along with more problem cases (additional single-instance cases and a larger database of design problems).
(2) A methodology for assessing the usefulness of solving design problems is introduced.
(3) A method for assessing the overall creative potential of design problems is proposed.
(4) Expert validation is added to ascertain the correctness and effectiveness of the proposed methods, thereby making them more robust and scalable. Wherever necessary, we have included and appropriately cited some of our earlier work to improve the completeness and clarity of this paper.

This paper starts with a literature review on design creativity and creativity assessment methods in Sections 2.1 and 2.2, respectively. As this research primarily focuses on the SAPPhIRE model of causality, Section 2.3 explains the SAPPhIRE model, and some observations from the literature are presented in Section 2.4. The research aim and question are explained in Section 3. Research methodology for assessment of creative potential via novelty and usefulness, their results and validation have been explained in Section 4. Section 5 presents the summary, conclusions and future scope, followed by the acknowledgements, references and appendix.

2. Literature review

2.1 Design creativity and problems

Creativity is the production of novel and useful ideas in any domain, and innovation is the successful implementation of creative ideas within an organization (Amabile 1996). Kao (1996) defines creativity as the process through which ideas are generated, developed and transformed into value. Creativity in engineering design occurs through a process by which a person uses their ability to generate ideas, solutions or products that are novel and useful (appropriate, social value) (Sarkar & Chakrabarti 2007). The definitions provided by Stein (1953), Amabile (1996) and Sarkar & Chakrabarti (2007), (2015), after their thorough analysis, highlight novelty and usefulness as the two prime indicators of creativity.

Creativity plays a crucial role in design problem formulation and solving, which is characterized by ill-structured and non-routine problems (Casakin & Kreitler 2005). The ability to restructure problems and develop new problem descriptions is essential for creative design solutions in both architecture and engineering (Akin 2008). Some problems, when properly structured, allow more space for exploration and thus foster creativity, whereas the narrower problems offer limited scope for substantial improvement (Goel & Pirolli 1992). An attempt has been made in Singh & Chakrabarti (2024) to assess the structure of the design problem with the intention that it may lead to an understanding of the creativity of design problems.

While creativity in design encompasses both problem-exploring and problem-solving aspects, traditional focus has largely been on problem-solving, with very

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

limited attention given to problem-exploring (Nidamarthi *et al.* 1997; Obieke, Milisavljevic-Syed & Han 2020).

To address this imbalance, emergent technologies such as data mining, natural language processing and machine learning can support human efforts in identifying creative design problems, potentially leading to breakthroughs in global problem exploration and inspiring more innovative solutions (Obieke *et al.* 2020, 2023). In architectural design studios, key factors influencing creativity include pedagogy, cognitive approach, interaction, information representation and creativity measurement (Casakin & Wodehouse 2021). Computational tools can enhance creativity in design, particularly for form finding, complexity optimization and ideation variety (Hanna 2018). In urban design, generative methods utilizing artificial intelligence (AI) have emerged to efficiently explore complex solution spaces and generate design options that satisfy conflicting objectives (Jiang *et al.* 2023). Although these approaches overcome the limitations of traditional manual design processes, in engineering, AI is not capable of combining elements to create something novel, which is the space the human still necessarily needs to fill (Lockhart 2025). Human creativity involves unique neurobiological machinery and experiential components that AI cannot replicate (Aru 2025). Hendler (2006) argues that AI excels at well-defined problems but struggles to combine known elements in novel ways, emphasizing the need to explore human cognitive capabilities that remain beyond AI's reach. These studies collectively highlight the importance of creativity in various design disciplines and the potential for computational tools, along with human efforts, to enhance creative problem formulation and solving.

2.2 Creativity assessment methods

Past research has focused on developing systematic approaches to assess creativity in product design, addressing both the design process and outcomes. Feldhusen & Goh (1995) assessed creativity comprehensively, considering multiple measures of cognitive processes, motivations and environmental factors. Moss (1966) defined creativity as a combination of usefulness and unusualness. Unusualness is rated on a 0–3 scale based on how different a solution is from existing ones. Usefulness is also scored 0–3 by comparing it with a benchmark solution that meets all requirements. Shah, Smith & Vargas-Hernandez (2003) developed metrics to evaluate the effectiveness of idea generation methods based on the quantity, quality, novelty and variety of ideas generated, proposing two experimental approaches to evaluate the effectiveness of idea generation methods: a direct method that examines the idea as a whole and an indirect method that decomposes the idea into its components and studies their individual and interactive effects. Ranjan, Siddharth & Chakrabarti (2018) proposed a method using novelty and requirement satisfaction as key indicators of creativity throughout the design process. Yuan & Lee (2013, 2014) developed a computational approach that quantifies the design process and correlates it with creativity ratings, establishing a formula for creativity assessment. Their method identified six critical factors in the design process that strongly correlate with creative outcomes. Yin *et al.* (2021) compared four creativity assessment methods, finding that the Consensual Assessment Technique (CAT) and Creative Product Semantic Scale (CPSS) produced consistent rankings between experts and non-experts. These studies aim to

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

improve the validity and reliability of creativity assessment in product design, offering potential applications in both academic and industrial settings. Chakrabarti & Khadilkar (2003) measured product novelty by comparing it with a reference product. They used vertical criteria (like need and technology) and horizontal criteria (main and supplementary functions) with differences weighted and summed to calculate novelty.

Lozano (2009) described creativity as innovation degree (novelty) and requirement accomplishment (usefulness). Novelty is rated as non-innovative, moderate or radical, and usefulness uses weights based on requirement types. These weights are often unknown during design, and usefulness scores may lack detail on which requirements are met. Sarkar & Chakrabarti (2007), (2011) combined assessment scores of novelty and usefulness to compute scores for creativity. Novelty assessment is carried out using SAPPhIRE models to compare products at different abstraction levels, as shown in Figure 1. Usefulness is assessed as a product of the following factors: (L) level of importance (LoI), (R) rate of the popularity of use, (F) frequency of usage and (D) duration of use or duration of benefit per usage. The LoI is assigned on a scale of 1–5 according to the purpose of the product being designed. The other factors, R, F and D, apply only to those products that are already out in the market. Sternberg & Lubart (1999) define "usefulness" in terms of "appropriateness." Similarly, Mumford & Gardner (1994) define "utility" in terms of "usefulness, appropriateness and social value." From these definitions, we understand that "useful" things are those that carry "social value." Sarkar & Chakrabarti (2011) further argued that the usefulness of a product is indicated by its use (i.e., a product is not useful if it is not used) and proposed the above four parameters for the assessment of usefulness of a product.

Srinivasan & Chakrabarti (2010b) used a similar novelty assessment, comparing new ideas to those in the existing idea space. The distinction is that the SAPPhIRE model's various abstraction levels assign scores of $n_a = 7$, $n_s = 6$, $n_i = 5$, $n_{ph} = 4$, $n_e = 3$, $n_r = 2$ and $n_p = 1$; if there is no difference at any of those abstraction levels, $n = 0$.

2.3 SAPPhIRE model of causality

The SAPPhIRE model of causality was developed by Chakrabarti *et al.* (2005), as shown in Figure 2, to explain the causality of natural and engineered systems and is often cited for novelty assessment. The model gets its name from the highlighted letters of its constructs: Action, State change, Physical Phenomena, Physical Effect, Input, oRgan and Parts. They together cover the system's physical components, interfaces, interactions, structural context and scientific laws governing them and provide a rich, comprehensive description of the functioning of biological and technical systems. To improve the comprehension of the constructs and, consequently, the application of the model, Srinivasan & Chakrabarti (2009) improved the clarity of the definitions of these constructs. Later, using appropriate rules and reasoning, Bhattacharya *et al.* (2024) improved the process for automated extraction of SAPPhIRE abstraction levels from a natural language document. Singh & Chakrabarti (2025) interpreted the three types of SAPPhIREs – normal state, failure state and recovered state proposed by Siddharth, Chakrabarti & Ranganath (2020) as solution, problem and intended solution SAPPhIREs, respectively, and

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

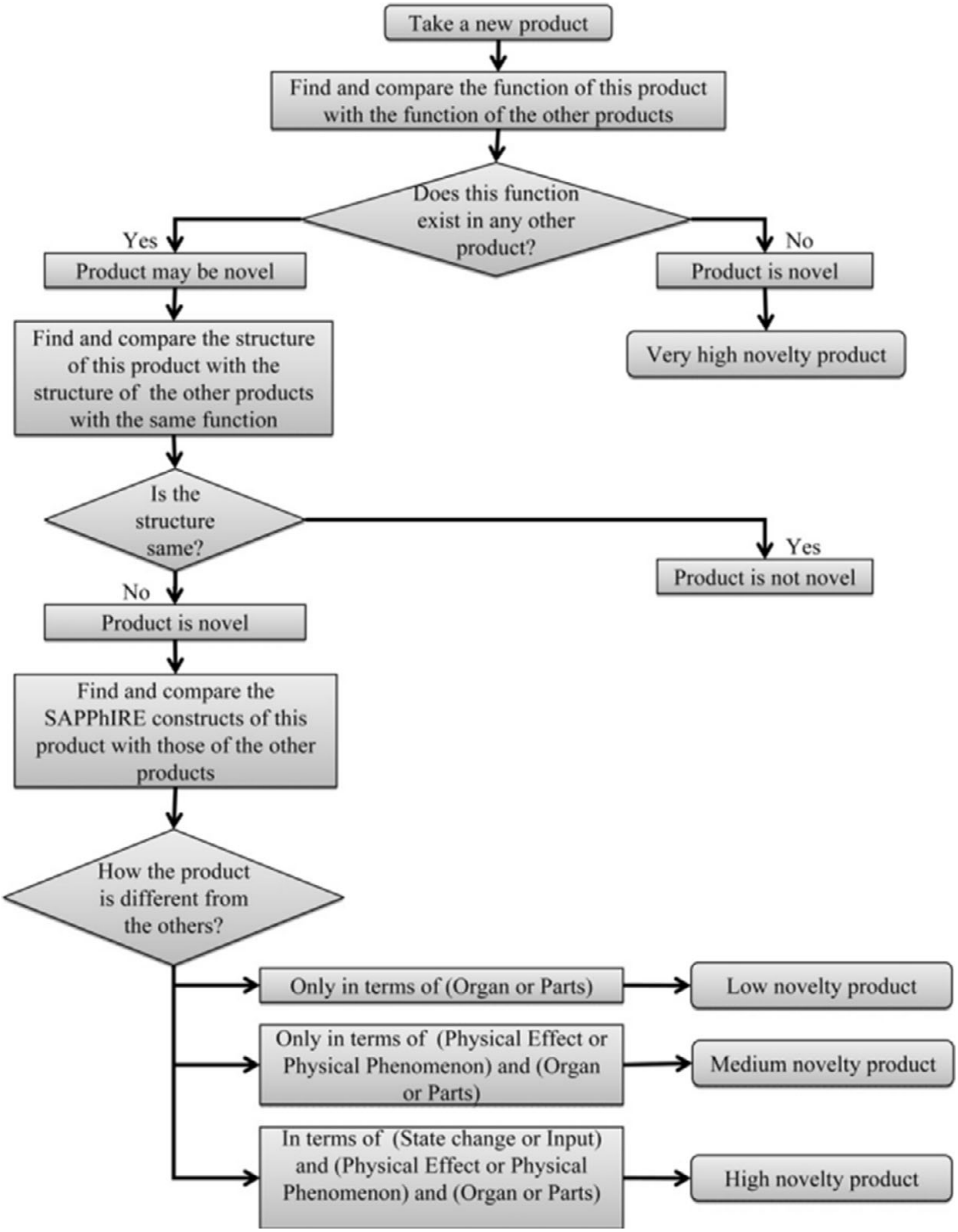


Figure 1. Proposed novelty assessment method by Sarkar & Chakrabarti (2007).

used problem SAPPhIREs in their preliminary framework for assessment of problem novelty.

The comparison between abstraction levels for reference and current products in terms of their novelty was illustrated by Srinivasan & Chakrabarti (2010b) and Ranjan *et al.* (2018). As a method for usefulness, Ranjan *et al.* (2018) evaluated the degree of requirement satisfaction (DRS) against solutions using the weighted objective tree method. Maher & Fisher (2012) assessed novelty, value and surprise

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

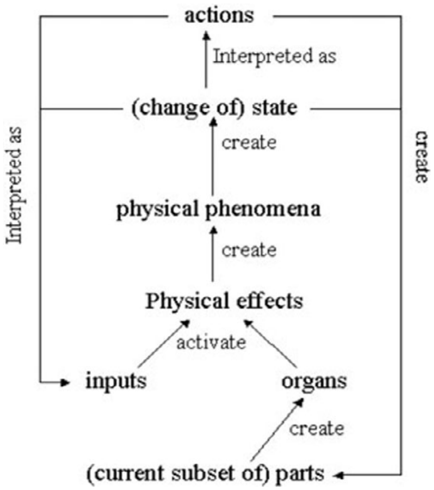


Figure 2. SAPPhIRE model of causality (Chakrabarti *et al.* 2005).

by comparing products in the same domain, where products are represented as attribute vectors and novelty is represented as the Euclidean distance between them. Grace *et al.* (2015) identified novelty and value as key features of creative design, wherein novelty is measured by the distance from a design to others using clustering and normalization; and value is estimated by the device's utility or popularity, such as web search results. The distance between word vectors as a method for comparing SAPPhIREs for idea novelty was proposed by Siddharth, Madhusudanan & Chakrabarti (2019).

2.4 Observations from the literature

Design problem identification and formulation have a strong influence on the creativity of solutions it produces; even a slight change in formulation can lead to very different design directions for the solution. By examining how the problem is stated, we can anticipate whether it is likely to spark novel and useful ideas (Wang *et al.* 2015). The novelty of design problems is defined as the extent to which they are new and original, serving as a direct translation of the definition originally applied to design solutions. For instance, the creation of Airbnb, which is unquestionably novel, was a novel solution to the novel problem of affordable and distinctive lodging options, particularly for tourists looking for experiences outside of traditional hotels (Guttentag 2013). In contrast to novelty, the definition of usefulness cannot be directly translated when applied to design problems, as opposed to design solutions. The usefulness of solving a design problem indicates only the potential usefulness to society if the design problem is solved well. In contrast, the usefulness of the solution represents the realized usefulness when the problem is worth solving and is solved well. We argue that if the problem is not solved well, the solution will not be a useful one. However, if the problem lacks potential usefulness, no matter how well it is solved, the solution will not be useful. If we choose a problem arbitrarily, potential usefulness is left to the mercy of chance, regardless of how well all its alternative solutions solve it. Hence, the

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

importance of usefulness of a design problem is the potential usefulness (scope of usefulness) created by the problem. Choosing an arbitrary solution at the end of the problem-solving process does not necessarily ensure that a problem's potential is realized.

Creativity of design solutions is defined as the product of novelty and usefulness (Sarkar & Chakrabarti 2011), and this formulation may also be extended to define the creative potential of design problems. The difference between problem creativity and solution creativity is that in problem creativity, among possible, alternative problems that are explored, some provide greater creative potential: are more novel than others (i.e., more different than the problems that have existed up until now) and are potentially more useful to solve than others. For any of the above problem alternatives, each with a different creative potential, there can be multiple possible solutions, some of which would be better at achieving the creative potential offered by the problem than others. Problem creativity is about creating the potential for being creative, while solution creativity is about realizing the potential. If the potential to start with does not exist or is limited during problem identification, it limits the scope of what can be realized during problem-solving. Assessing the creative potential of design problems adds an additional layer to compound creativity, as the current focus remains limited to the creativity of design solutions alone.

The methods and metrics have been proposed for the design outcomes, such as products or solutions, but not for the creativity of design problems, which is a bigger research gap, as found out. There are very few qualitative approaches in the literature that have examined this topic. A simple method for accomplishing this has been proposed by Roser & Dartnall (2007). It incorporates the use of simple assessment indicators, rating the level of problem novelty and complexity. Since these qualitative measurements can only provide subjective judgments based upon problem understanding, a quantifiable metric is required to overcome this bias. Singh & Chakrabarti (2025) made one of the initial attempts to quantitatively assess the novelty of design problems with a limited and simpler problem set and analysis. Current literature also lacks in the assessment of the (potential) usefulness and creative potential of design problems. The study proposed by Sarkar & Chakrabarti (2011) incorporates the quantitative assessment of the usefulness and creativity of design solutions. However, the parameters proposed in this study could also be used as indicators for whether a design problem, if solved, is likely to be useful or not.

3. Research aim and questions

Based on the observations drawn from the literature review, we have determined the following as the aim and objectives for this research:

- It is necessary to assess the creative potential of design problems to identify those problems that are both more novel and more valuable (and hence more creative).
- To do this, a generic method for identifying, analyzing and evaluating design problems for their creative potential is needed, which could be used for various stakeholders in a variety of systems and contexts.
- Singh & Chakrabarti (2025) proposed a preliminary methodology for the assessment of the novelty of design problems; however, more analysis and validation against expert judgment are required to better support their evaluation.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

To address this aim, the following research question has been framed, with three sub-questions:

1. How to assess the creative potential of a design problem?
   1.1 How to assess the novelty of design problems?
   1.2 How to assess the usefulness of solving design problems?
   1.3 How to assess the creativity of design problems?

4. Research methodology

4.1 Adaptation of the method for assessment of novelty of a design problem

An evaluation of the existing methods for assessing novelty (Srinivasan & Chakrabarti 2010b; Ranjan *et al.* 2018) reveals the following: For comparing solutions, these methods use a scoring system that gives higher values to ideas or solutions that differ at higher abstraction levels than those at lower abstraction levels in the SAPPhIRE model of causality. However, various aspects of a problem may exist at different levels of abstraction. Therefore, a problem cannot necessarily be assessed at a single level of abstraction; each aspect of the problem may have to be addressed at its relevant level of abstraction. Assessment for design problems with a single cause (represented using single-instance SAPPhIREs) is demonstrated in our previous study (Singh & Chakrabarti 2025). For complex problems with multiple root causes, the problems need to be represented using a multi-instance problem SAPPhIRE model and subsequently assessed. Based on this understanding, the previously developed method by Singh & Chakrabarti (2025) is adapted for a broader set of problems and for a detailed evaluation of novelty of design problems. Method steps are explained below.

To address the first research sub-question (how to assess the novelty of a design problem), data on similar design problems are collected from primary stakeholders, i.e., designers (of 2–3 years' experience in the product design domain) and customers (using such products for 3 years or more), for their design solutions. The research methodology (Figure 3) involves the following steps:

1. Existing design problems in the same context from patents or online searches are collected, and their classification as "past problems" is carried out.
2. Current (i.e., new) design problems are collected using a questionnaire and survey, as faced by stakeholders (designers and customers) in that context, and classified as "current problems."
3. Each of the past and current problems so collected is described using the SAPPhIRE model of causality, henceforth called problem SAPPhIRE (Siddharth *et al.* 2020). The completeness of SAPPhIRE models is taken care of using the method given by Bhattacharya *et al.* (2024). Design problems with a single root cause are represented using single-instance problem SAPPhIREs, while those with multiple root causes are represented using multi-instance problem SAPPhIREs. A multi-instance problem SAPPhIRE is constructed by linking multiple single instances, where the input of each next instance is derived from the action level of the previous instance.
4. The natural language description of each past and current problem SAPPhIRE in these two distinct databases is compared with one another for similarity at the action level of SAPPhIRE (the comparison is facilitated using Bidirectional

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

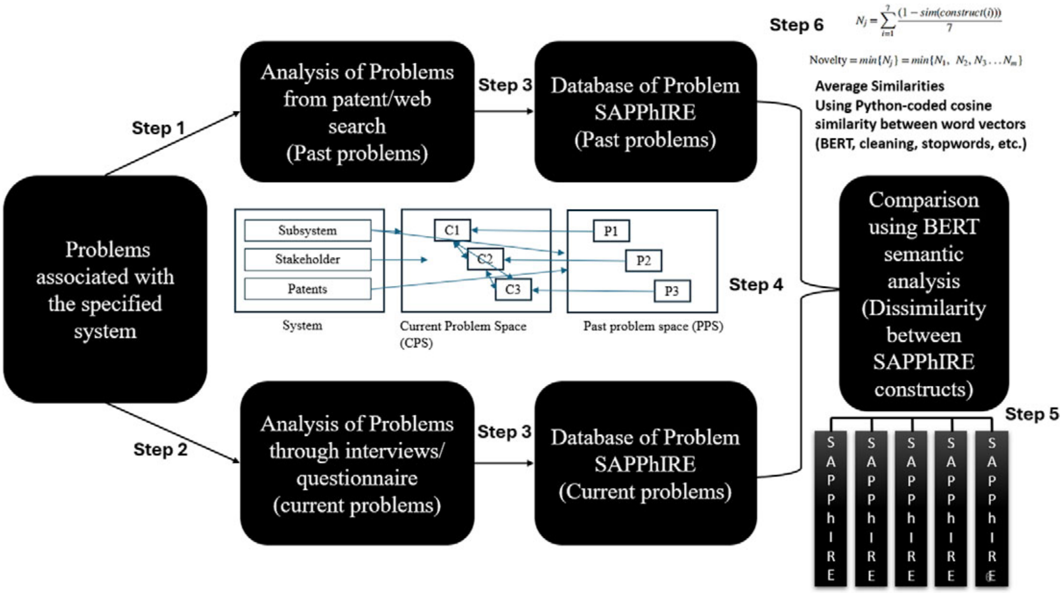


Figure 3. Proposed methodology for the assessment of the novelty of design problems (adapted from Singh & Chakrabarti 2025).

Encoder Representations from Transformers (BERT) semantic similarity with a pre-trained dataset). For multi-instance SAPPhIREs, comparison is carried out after concatenating the sentences at each abstraction level of SAPPhIRE.

5. For those problem SAPPhIREs for which the action level of SAPPhIRE matches, sentence similarity and word-to-word similarity are calculated between each pair of problem SAPPhIREs by converting their words into word vectors and measuring the cosine distance between them at each remaining level of SAPPhIRE abstraction (Siddharth *et al.* 2019). Semantic similarity is first assessed between the word vectors and then extended to the sentences. The pre-training of data is performed using the model "sentence-transformers/all-MiniLM-L6-v2."
6. Similarity score for each level of SAPPhIRE abstraction is then calculated, where dissimilarity (novelty) scores for each pairwise comparison are determined using the following equation: novelty = (1 – similarity) (Siddharth *et al.* 2019).
7. Overall novelty score of the given, current problem is now calculated, in comparison with the previous problems under consideration, by averaging these pairwise novelty scores.
8. The novelty score is interpreted qualitatively using the following ranges: 0 to 0.3 as low novelty; 0.3 to 0.7 as medium novelty; and 0.7 to 1 as high novelty.

Using the pre-trained dataset and a web-based tool for sentence comparison, some of the above processes (Steps 4 and 5) are automated. To further visualize the better applicability, a case study on one system is explained below.

4.1.1 Case study (electric kettle) demonstrating the application of the method

An electric kettle is used as the system for problem analysis in our preliminary case study (Singh & Chakrabarti 2025). Different parts of the electric kettle with their

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

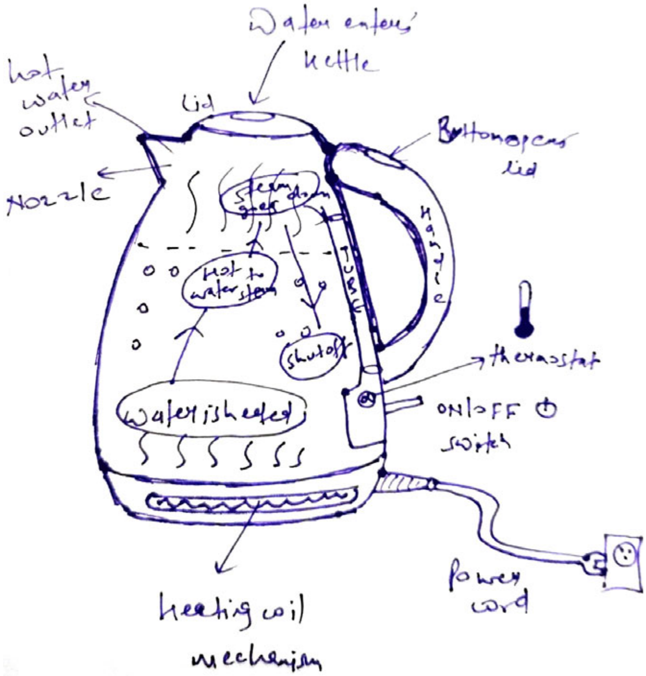


Figure 4. Electric kettle (Miller 2019).

functions are shown in Figure 4. Utilizing online patent libraries and patent searches, a comprehensive examination of patents pertaining to electric kettles from the last 100 years is carried out. Beyond the initial data collection, additional design problems are collected, encompassing both single-cause and multi-cause cases. After that, useful data in the form of past design problems are extracted, and using these, a database of past problems for this system is developed. In Table 1, a few of the problems collected are listed. We assume that the problems gathered from patents are "old" and problems gathered currently through a questionnaire are "new."

Table 1. Some of the design problems collected from patents and web sources

S. no.	Past design problems
1.	Avoiding overheating of the liquid present in the kettle
2.	Providing safety to the users from the hot steam coming out
3.	Reducing heat transfer out of the kettle
4.	Spilling of the liquid decreases the efficiency of the kettle
5.	Temperature and boiling control to minimize inefficient heating
6.	Tough to clean

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table 2. Some of the design problems collected from the survey questionnaire

S. no.	Current design problems
1.	Water is not getting heated
2.	Scalding of the human hand due to hot steam
3.	Overheating takes place since there is no temperature monitoring available
4.	Spilling of liquid due to liquid expansion
5.	Cleaning is tough inside the intricate areas of the kettle base
6.	Underheating of the coil element due to a corroded coil
7.	Localized roughness in the kettle base due to longer scrubbing results in local heating/boiling
8.	The conical shape makes it difficult to wash
9.	Non-optimal heating due to a faulty thermostat

Moving forward, we collected a set of current design problems, as mentioned by the stakeholders (i.e., designers and customers), using a survey questionnaire (see Appendix, Figure A1). These problems are stored in a database of current problems. Some of these problems are enlisted in Table 2.

The problems written in Tables 1 and 2 are ill-structured problems which, during the application of methodology, are converted into the unified category of "problems to be solved in the design process," but somehow, they are the problems that exist at the beginning of a design process. The past problems are analyzed and transformed into their problem SAPPhIREs using the problem SAPPhIRE model, as shown in Figure 5 (Siddharth *et al.* 2020) and rules given by Bhattacharya *et al.* (2024). The current problems (at the time of analysis) are also analyzed and transformed into their problem SAPPhIREs. Both are stored in the problem SAPPhIRE database (see Appendix, Table A1).

The problems of "spilling of liquid," "non-optimal heating," and "scalding of human hand" are selected as the example problems for novelty assessment, as they belong to different categories, i.e., efficiency and safety, and also based on their high frequency of occurrences as reference past problems for this case (see Appendix, Table A1). Similarly, new problem SAPPhIREs are chosen as current problems from the database. The design problem is the opposite of the action achieved by the solution (Siddharth *et al.* 2020). So, the action level, which is the highest level of abstraction, is used as a design problem in this case (Figure 5). There may be instances where the problem exists at different levels of abstraction that authors have not yet encountered. So, for comparison, their action-level description includes "spilling of liquid." The database (see Appendix, Table A1) is managed using a Microsoft Excel tool created by the researcher to save time while creating problem SAPPhIREs.

A search for common, action-level SAPPhIREs is needed, since only similar problems can be compared for novelty assessment. While selecting SAPPhIREs for problems, there may be situations where a single instance is present, while others have multiple instances. A single instance indicates that the structure includes one set of SAPPhIRE elements, while a multi-instance structure contains multiple, connected sets of SAPPhIRE elements, typically due to the presence of several root causes. Direct comparison is used for single instances (Figure 6). The abstraction

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

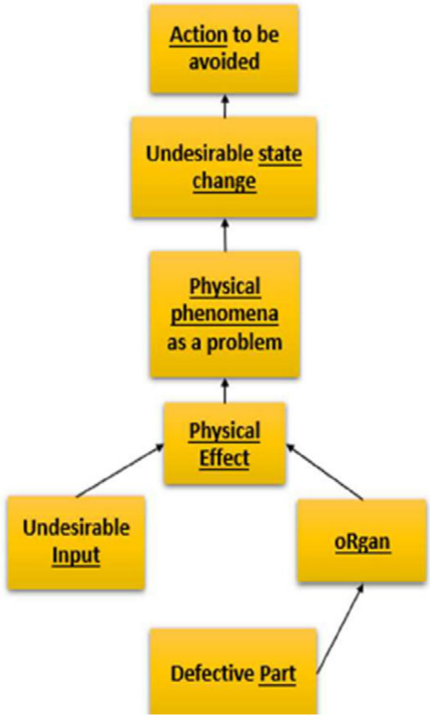


Figure 5. Problem SAPPhIRE model (Siddharth *et al.* 2020).

levels are first concatenated in the case of multiple instances, and then, the same comparison approach is used, as illustrated in Figure 7. Following a comparison of sentence semantics at the action level, the semantic similarity of sentences is compared at further abstraction levels, which include state change and phenomena. Some of the formulated and assessed design problems, drawn from the broader set of assessed problems (see Appendix A1, Table A2), are presented below:

1 Spilling of liquid due to liquid expansion
2 Spilling of liquid due to seal and gasket rupture
3 Non-optimal heating due to a faulty switch and damaged wire
4 Non-optimal heating due to a faulty thermostat
5 Non-optimal heating due to reduced contact area
6 Scalding of the human hand due to hot steam

4.1.2 Results

The comparison is carried out among each set of problems, using their current and past problem SAPPhIREs. Comparison scores corresponding to each abstraction levels are computed and integrated into an overall score for the entire problem. These are tabulated for some of the problems in Column 3 of Table 3. Based on the proposed range of quantitative values for qualitative interpretation (see Section 4.1), inferences are made about the degree of novelty for each problem (see Column 4 of Table 3). The level of novelty (high, medium and low) of the design

Figure 6. Comparison between two single-instance problem SAPPhIREs.

Action: No snap of thermostat. Non-closing of switch. Non-optimal heating *State Change*: Curved to unflexed position. Switch on to on position . Change in Coil temperature *Phenomena:* Heat energy converts into strip snap action. Less torque will not trip the switch. High heat generation *Physical Effect:* Expansion equation LαdT. F x L = Torque. Joule's law of heating *Input:* Hot steam. Less force to lever. High electric current flow *oRgan:* Coefficient of thermal expansion. Lever length. Coil resistance *Parts:* Faulty Bi-metallic Thermostats. Lever. switch	*Action:* Uneven rough surface. Non-optimal heating *State Change*: Smooth to rough surface. Normal to less heated area *Phenomena:* Excessive scrubbing leads to abrasion of material. Reduced contact area leads to uneven heating *Physical Effect:* Principle of abrasion. Conduction law *Input:* Excessive scrubbing. Heat input *oRgan:* Base area, base material. Reduced contact area *Parts:* Base of kettle. Kettle base

Figure 7. Comparison between two multi-instance problem SAPPhIREs after concatenation of sentences.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table 3. Novelty scores as computed from the proposed method for different design problems

S. no.	Problem's description	Novelty scores (least) from proposed assessment	Cumulative decision on novelty
1.	Spilling of liquid due to liquid expansion	0.55	Medium novel
2.	Non-optimal heating due to a faulty thermostat	0.74	High novel
3.	Scalding of the human hand due to hot steam	0.27	Low novel
4.	Spilling of liquid due to seal and gasket rupture	0.65	Medium novel
5.	Non-optimal heating due to reduced contact area	0.8	High novel
6.	Non-optimal heating due to a faulty switch and damaged wire	0.67	Medium novel

problems related to an electric kettle, according to their scores, is shown in Table 3 (5 having the highest novelty and 3 having the lowest):

4.1.3 Expert validation for correctness and effectiveness of the proposed method

In design literature, several researchers have emphasized expert judgment as one of the most reliable approaches for evaluating the creativity of design solutions (Amabile 1996; Sarkar & Chakrabarti 2011). Accordingly, for the correctness of the proposed method, the inferences that are made for the degree of novelty of assessed design problems need to be compared with the intuitive notion of experts.

Three experts are chosen, and they are asked to assess the novelty of a representative set of design problems, selected to capture a wide variety while minimizing the resource-intensive (time and effort-intensive) nature of evaluation. Each expert had at least three years of experience in new product development and is aware of the SAPPhIRE model of causality. To ensure a diverse perspective, the experts are from different regions of the world, accounting for diversity in product understanding and usage.

As illustrated in Figure 8, the experts evaluated six design problems ((1) spilling of liquid due to liquid expansion; (2) non-optimal heating due to a faulty thermostat; (3) scalding of the human hand due to hot steam; (4) spilling of liquid due to seal and gasket rupture; (5) non-optimal heating due to reduced contact area; and (6) non-optimal heating due to a faulty switch and damaged wire) from a similar context by rating their novelty at five abstraction levels (which are expressed in more general terms to make it suitable for the experts), consolidated from the original seven for greater clarity (Siddharth *et al.* 2019).

The general terms (see Appendix, Table A2 for definitions) are purpose (Action), output (State Change), process (Physical Phenomena and Effect), input and components (Parts and oRgan). Novelty scores ranged from low (0.3) to medium (0.6) to high (0.9). Novelty scores given by all three experts, abstraction-level-wise, are

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Expert	Abstraction level	P1- Spilling of liquid in liquid heating device	Novelty Scores	Abstraction level	P2- Non optimal heating in liquid heating device	Novelty Scores	Abstraction level	P3- Scalding of human hand in liquid heating device	Novelty Scores
E1	Purpose	Spilling of liquid	0.3	Purpose	non optimal heating	0.6	Purpose	scalding of human hand	0.6
	Output	Static to movable liquid	0.3	Output	no change in coil temperature due to switching on of switch caused by unflexed thermostat	0.6	Output	excess steam leads to burnt hand	0.3
	Process	Overheating the liquid leads to its expansion using conduction law	0.6	Process	high heat generation due to non closing of switch provided by less torque from lever caused by non snap action of thermostat	0.6	Process	hot steam due to heat transfer start burning the hand	0.3
	Input	High heat energy in the whole tank volume	0.6	Input	high electric current due to less force from lever and hot steam	0.6	Input	uncontrolled heating	0.6
	Components	Heater tank	0.3	Components	switch , lever, faulty thermostat	0.6	Components	kettle, human hand	0.3

Figure 8. Template for expert scoring based on their intuition.

Expert	Abstraction level	P1	P2	P3	P4	P5	P6
E1	*Purpose(Action)*	0.3	0.6	0.6	0.3	0.6	0.6
	Output(State Change)	0.3	0.6	0.3	0.6	0.6	0.3
	Process (Phenomena & Effect)	0.6	0.6	0.3	0.6	0.9	0.6
	Input	0.6	0.6	0.6	0.3	0.9	0.3
	Components(Parts & oRgan)	0.3	0.6	0.3	0.6	0.6	0.3
E2	*Purpose(Action)*	0.9	0.6	0.3	0.6	0.6	0.6
	Output(State Change)	0.9	0.6	0.3	0.6	0.9	0.3
	Process (Phenomena & Effect)	0.3	0.9	0.3	0.3	0.3	0.3
	Input	0.3	0.3	0.3	0.3	0.3	0.3
	Components(Parts & oRgan)	0.3	0.3	0.3	0.3	0.3	0.3
E3	*Purpose(Action)*	0.6	0.6	0.3	0.6	0.6	0.6
	Output(State Change)	0.6	0.6	0.6	0.6	0.9	0.6
	Process (Phenomena & Effect)	0.3	0.6	0.3	0.6	0.9	0.6
	Input	0.9	0.6	0.3	0.9	0.9	0.6
	Components(Parts & oRgan)	0.6	0.9	0.3	0.6	0.9	0.6

Figure 9. Novelty scores by three experts for all five abstraction levels across six different design problems.

shown in Figure 9, with the values distinguished by variation in color. Dark green shows highly novel, light green followed by yellow for medium novelty, and red shows low novelty. Judgment is based upon the experience and intuition of experts; they were not asked to follow any method.

Considering the degree of novelty of the whole problem (low – L, medium – M and high – H; explained in Section 4.1), in four out of the six cases, all three experts agree with the researcher's scores and inferences as obtained from the proposed assessment of design problems (see Table 4). In five of the six cases, at least one expert out of the three agrees with the researcher's scores and inferences. Another way is to count all the matches (Cohen 1960) and divide them by the total number of pairs compared (see Table 4). Out of 18 pairs compared, the number of matches

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table 4. Comparison of the researcher's novelty scores with those of experts for various design problems

Scorer/ problems	P1 Spilling of liquid due to liquid expansion	P2 Non-optimal heating due to a faulty thermostat	P3 Scalding of the human hand due to hot steam	P4 Spilling of liquid due to seal and gasket rupture	P5 Non-optimal heating due to reduced contact area	P6 Non-optimal heating due to a faulty switch and damaged wire
Researcher	0.55 (M)	0.74 (H)	0.27 (L)	0.65 (M)	0.8 (H)	0.68 (M)
Expert 1	0.42 (M)	0.6 (M)	0.42 (M)	0.48 (M)	0.72 (H)	0.42 (M)
Expert 2	0.54 (M)	0.54 (M)	0.3 (L)	0.42 (M)	0.48 (M)	0.36 (M)
Expert 3	0.6 (M)	0.66 (M)	0.36 (M)	0.66 (M)	0.84 (H)	0.6 (M)

is (3 + 0 + 1 + 3 + 2 + 3 = 12); so % match is 12/18 = 0.67. This indicates that there is over 67% agreement with the scores of all the experts and over 83% agreement with the scores of at least one expert. Abstraction-level-wise agreements are shown in Figures 10–12.

Problem 5 has received high novelty rankings, and problems 1, 4 and 6 have received medium novelty rankings from all three experts, with which the researcher also agrees. Figure 11 compares the novelty scores given by researchers and experts for the state change level. It is noted that the state change level, due to its higher level of abstraction, provides a higher novelty score. A similar trend exists at the phenomenon and effect levels. This is because more problems are typically found at the part or organ level, and so the similarity, where the experts have given a lower novelty

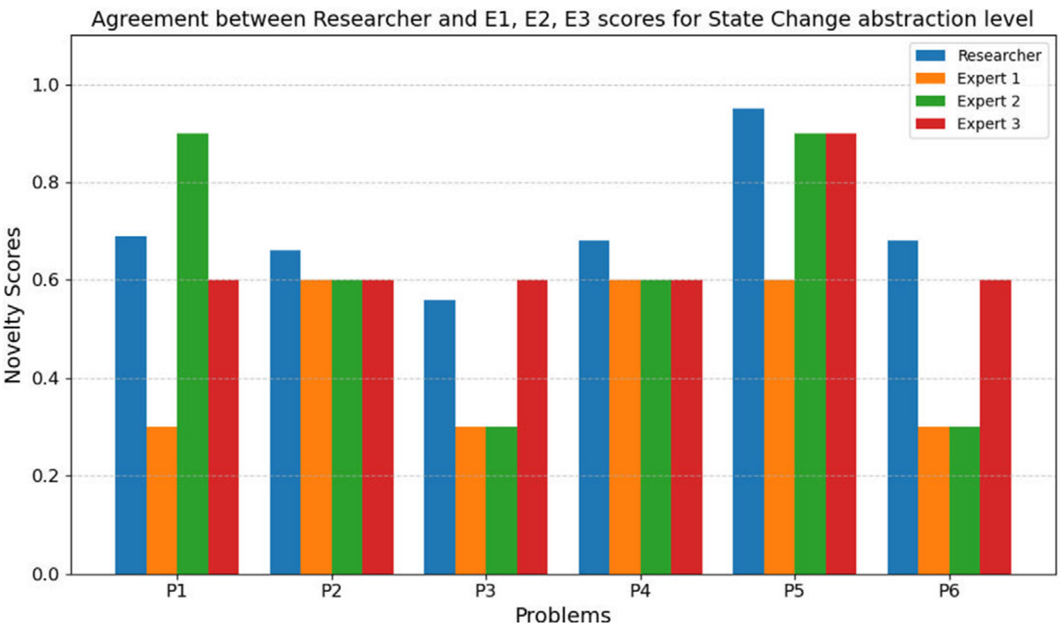


Figure 10. Comparison between the researcher's and experts' scores for the state change level.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

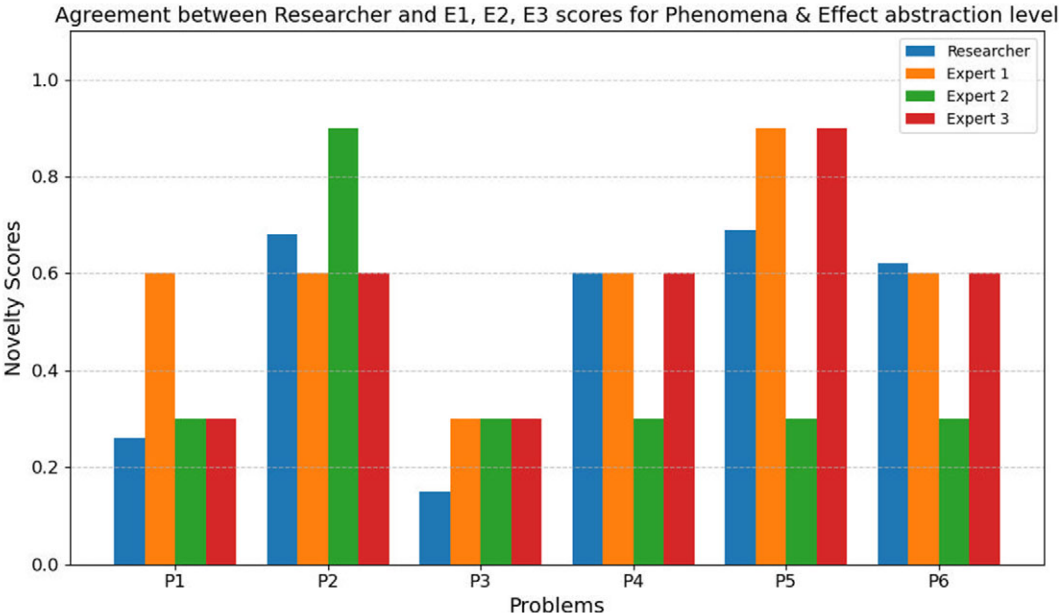


Figure 11. Comparison between the researcher's and experts' scores for phenomena and effect level.

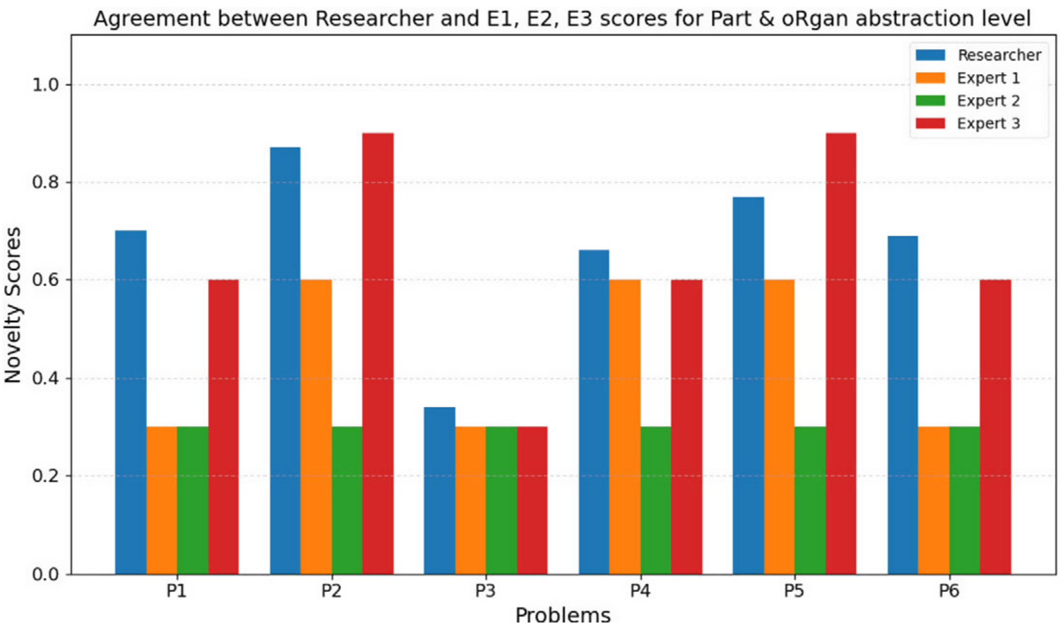


Figure 12. Comparison between the researcher's and experts' scores for the part and organ level.

score except for the case where a completely different part is invented. Problems at the phenomenon level are uncommon, like state change. According to Srinivasan & Chakrabarti (2010b), there is a strong correlation between novelty scores and levels of abstraction, which is partially supported by our research.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Furthermore, a comparison of the proposed method with one of the methods available in the literature (Srinivasan & Chakrabarti 2009) is carried out to evaluate the effectiveness of the proposed method. The method proposed by Srinivasan & Chakrabarti (2009) is a manual assessment method used for assessing the novelty of an idea. Novelty of an idea is calculated by looking at the highest level of abstraction of dissimilarity in the SAPPhIRE model of causality and assigning the scores based on the following ranges:

"If action level is dissimilar, assign a score of 7; if state change level is dissimilar, assign a score of 6; if input level is dissimilar, assign a score of 5; if phenomena level is dissimilar, assign a score of 4; if effect level is dissimilar, assign a score of 3; if organ level is dissimilar, assign a score of 2; if part level is dissimilar, assign a score of 1 and if nothing is dissimilar, assign a score of 0." After the score is assigned, it will be averaged considering seven no. of elements, and a final averaged score is used for qualitative inferencing of the degree of novelty.

We have applied this procedure for manual assessment on the corresponding sets of problem SAPPhIREs, which we have used in our proposed method, and calculated the novelty scores, respectively (as shown in Column 3 of Table 5).

The novelty scores from the proposed method (method 1; see Column 2 in Table 5) and the manual method (method 2; see Column 3 in Table 5) are compared to the average novelty scores of all three experts. The root-mean-square (RMS) formula is used to calculate the deviation between method 1 and method 2 scores and the expert scores. The RMS deviation between method 1 scores and experts' average scores is 0.13, while the RMS deviation between method 2 scores and experts' average scores is 0.30. The lower the RMS, the better

Table 5. Comparison of effectiveness of the proposed and the manual method using experts' average scores

Problem's description	Novelty scores from the proposed method – method 1	Novelty scores from the manual method – method 2	Experts' average novelty scores	Difference between method 1 scores and expert average scores	Difference between method 2 scores and expert average scores
P1 – Spilling of liquid due to liquid expansion	0.55	0.85	0.52	0.03	0.33
P2 – Non-optimal heating due to a faulty thermostat	0.74	1	0.6	0.14	0.4
P3 – Scalding of the human hand due to hot steam	0.27	0.42	0.36	−0.09	0.06
P4 – Spilling of liquid due to seal and gasket rupture	0.65	0.57	0.52	0.13	0.05
P5 – Non-optimal heating due to reduced contact area	0.8	1	0.68	0.12	0.32
P6 – Non-optimal heating due to a faulty switch and damaged wire	0.68	0.85	0.46	0.22	0.39

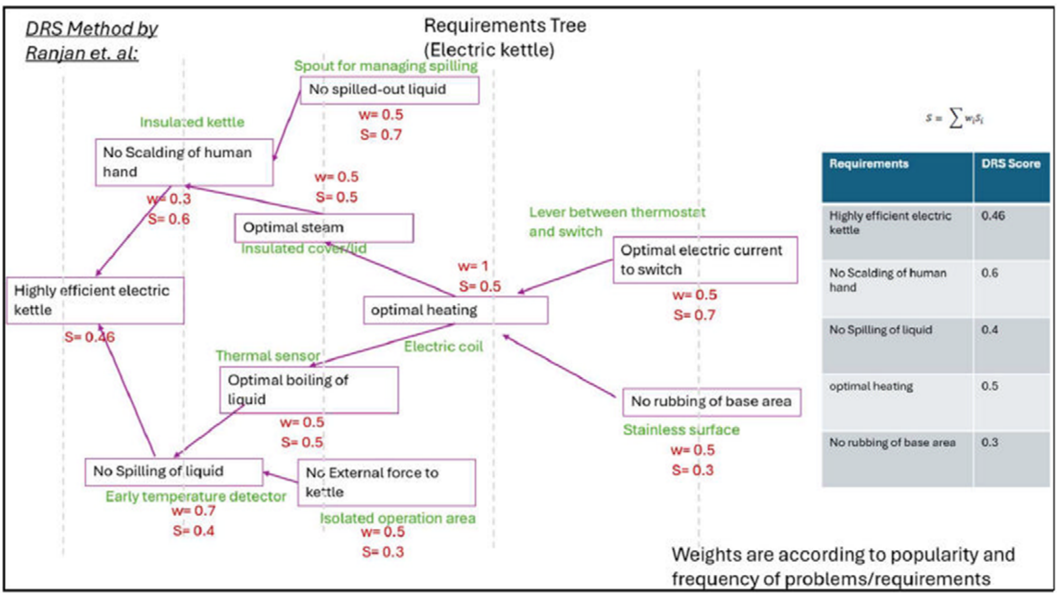
https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

the method is. This demonstrates that the proposed method is the least deviant from the experts' intuitive notion and therefore more effective than the existing method in capturing the intuitive notion of the experts. We argue that subjective models are developed primarily due to the limited availability of experts; however, when validating such models, expert judgment carries greater weight than the models themselves (Sarkar & Chakrabarti 2011).

4.2 Methodology for assessment of usefulness (of solving a design problem)

The weighted objectives method, which was adapted by Ranjan *et al.* (2018) for assessing "degree of requirement satisfaction (DRS)" as a proxy for assessing usefulness, and the parameter-based usefulness assessment method by Sarkar & Chakrabarti (2011) are two major, existing methods for assessing the usefulness of design solutions. We have employed the weighted objectives method adapted by Ranjan *et al.* (2018) as part of our earlier descriptive study (Blessing, Chakrabarti & Wallace 1995). In this method, low-level requirements are initially assigned weightages based on prior knowledge of users and requirement-satisfaction scores. Then, the weightages and requirement-satisfaction scores are calculated using the formula provided in Figure 13 for the next high-level requirements and so on. Following the same approach, we have determined the usefulness scores of design problems utilizing a problem-requirement tree as illustrated in Figure 13.

However, the method proposed by Ranjan *et al.* (2018) involves a higher degree of subjectivity, which limits its applicability. This subjectivity arises primarily from the distribution of weightages based on the prior knowledge of users, as well as the dependency of high-level problems on low-level problems for the calculation of requirement-satisfaction scores.



Requirements	DRS Score
Highly efficient electric kettle	0.46
No Scalding of human hand	0.6
No Spilling of liquid	0.4
optimal heating	0.5
No rubbing of base area	0.3

Figure 13. Weighted objectives method (Ranjan *et al.*, 2018) illustrating the calculation of DRS.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

To address the second research sub-question, that is, how to assess the usefulness of solving design problems, the method proposed by Sarkar & Chakrabarti (2011) is adapted instead because it is more comprehensive and objective, making it more suitable for assessing the (potential) usefulness of design problems, using the following three parameters:

1. **Prevalence among users** – If a larger number of people face a problem, it is more useful to solve it as it benefits a greater number of users.
2. **Importance to society** – There is greater usefulness accrued from solving those problems that more significantly impact the users' lives than those that do not.
3. **Frequency of occurrence** – If a problem arises more repeatedly or across more instances (than another), solving it is more useful (than solving the other).

We have performed a similar type of data collection as in our novelty assessment. The following are the steps in the methodology proposed (see Figure 14):

1. Design problems within the same context are collected through case studies, system analysis and stakeholder interviews/questionnaires (see Appendix, Figure A1).
2. Problems and requirements are organized into a hierarchical tree structure using the data collected to understand the importance of a problem (Ranjan *et al.*, 2018).
3. Usefulness indicator analysis is conducted using three literature-derived and modified variables:
   - Prevalence (representing the rate of popularity of the problem)
   - LoI (representing how important a problem is for stakeholders)
   - Frequency of the problem (how frequently a problem occurs for stakeholders)
4. Prevalence is calculated using the formula "n/N," where n = number of individuals experiencing the problem and N = total number of participants in the study.
5. The LoI and frequency of the problem through binary scoring (low: 0 and high: 1) is assessed by taking qualitative data from multiple people, quantifying it and then averaging the maximum binary scoring for a single inference, e.g., if the maximum values correspond to the category "high," they are assigned a value of 1, and the average of these values is then computed.
6. The weighted average usefulness score (Eqn. 1) is calculated by assigning weights to each variable through the analytic hierarchy process (AHP), using

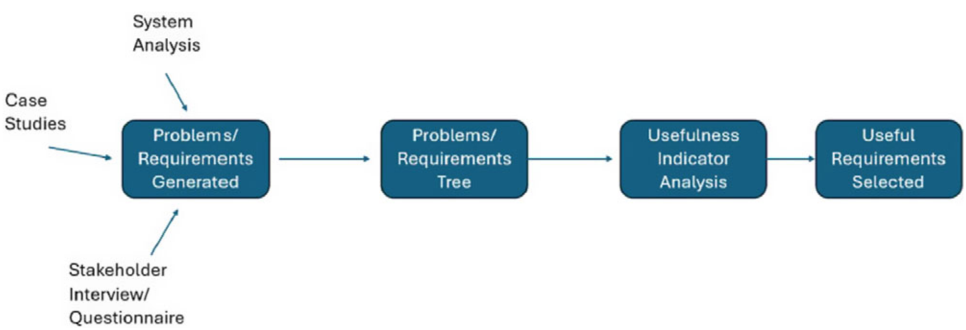


Figure 14. Proposed methodology for the assessment of usefulness of (solving) design problems.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

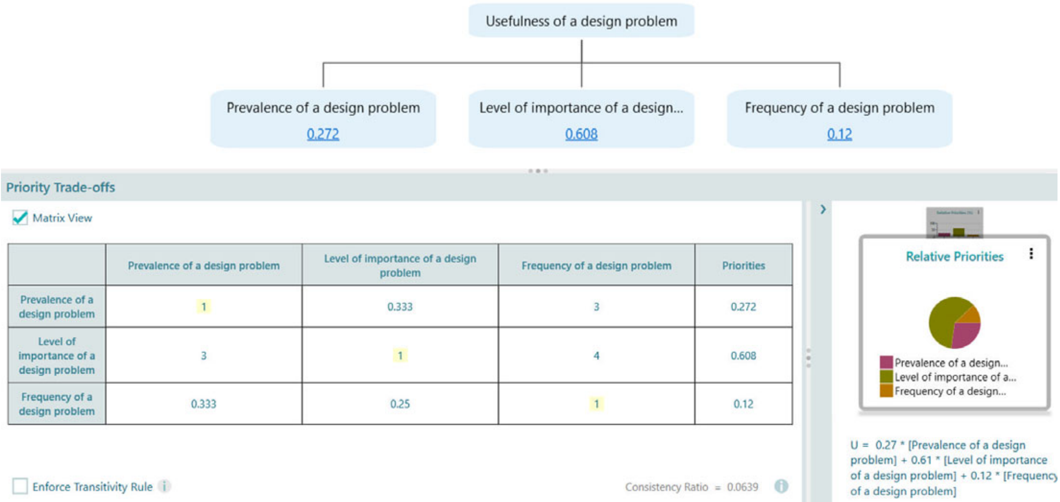


Figure 15. Weight distribution of the three variables for usefulness score using the analytic hierarchy process tool (https://www.spicelogic.com/Products/ahp-software-30).

an online tool to derive the weights (Figure 15), i.e., 0.27 for prevalence, 0.61 for LoI and 0.12 for frequency of a design problem.

$$\text{U (Usefulness score)} = 0.27\text{P} + 0.61\text{L} + 0.12\text{F} \tag{1}$$

where P is prevalence, L is LoI and F is frequency of a design problem.

7. Results based on the scores are interpreted: Higher values indicate more useful problems to address, while lower values suggest lower usefulness.
8. Usefulness of various design problems is assessed both within their original context and across different contexts, as shown in Table 6.

Table 6. Calculation of the usefulness score using values of all variables for different problems

Problems/ usefulness indicators	P = prevalence (n/N) (range 0–1) weight = 0.27	L = level of importance (low: 0 high: 1) weight = 0.61	F = Frequency of the problem (low: 0 high: 1) weight = 0.12	Weighted average score (U = usefulness score) U = 0.27P + 0.61 L + 0.12F
Problem 1				
Problem 2				
Problem 3				
Problem 4				
Problem 5				
Problem 6				

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

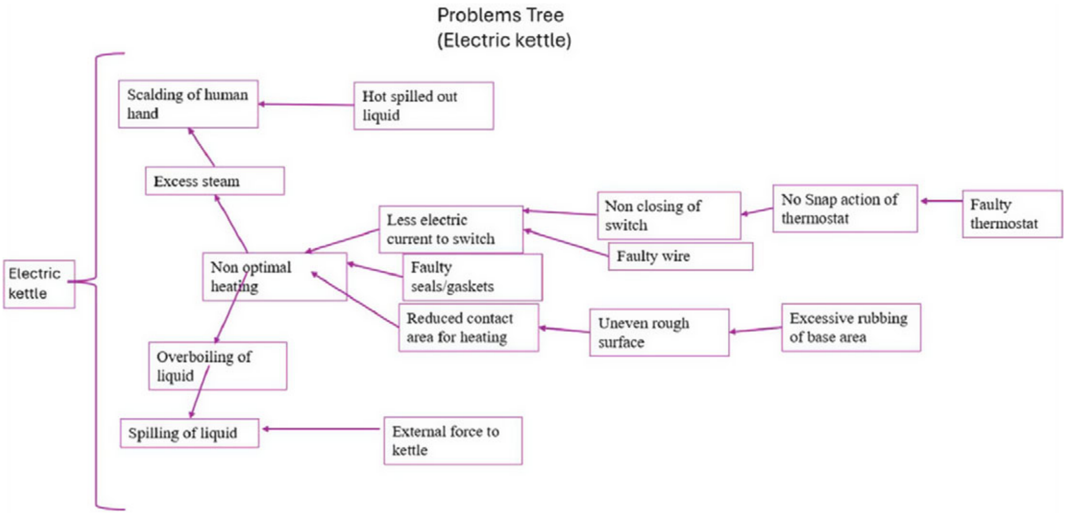


Figure 16. Problem tree for an electric kettle.

4.2.1 Case study (electric kettle) demonstrating the application of the method

A similar product to the one used in the novelty study is chosen for the demonstration of this method. Problems collected from the various sources are mapped into a tree as follows:

A high-level problem of "scalding of the human hand," and "spilling of liquid," followed by non-optimal heating, and subsequent low-level problems of "less electric current to switch," "reduced contact area of heating," "faulty thermostat," and so on, is shown in the problem tree in Figure 16.

The problems that are considered for assessment are as follows:

Problem 1: spilling of liquid due to liquid expansion; *Problem 2:* non-optimal heating due to a faulty thermostat; *Problem 3:* scalding of the human hand due to hot steam; *Problem 4:* non-optimal heating; *Problem 5:* non-optimal heating due to reduced contact area; and *Problem 6*: non-optimal heating due to a faulty switch and damaged wire.

4.2.2 Results

Twenty participants (customers and designers), who are the main stakeholders, have been given a questionnaire-based form (see Appendix, Table A3) with the following questions:

- Are you aware of this system and its problems?
- Are you encountering this problem?
- Do you feel it is important to solve this problem?
- How frequently do you encounter this problem?

To calculate the weighted average usefulness score, qualitative data pertaining to the necessary variables is recorded and transformed into a measurable metric, as shown in Table 7.

From the scores obtained, it is found that Problem 3 (scalding of the human hand due to hot steam) is the most useful to solve, while Problem 4 (spilling of

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table 7. Usefulness score with values of all the variables for different problems

S. no.	Problems/ usefulness indicators	P = prevalence (n/N) (range 0–1) weight = 0.27	L = level of importance (low: 0 high: 1) weight = 0.61	F = frequency of the problem (low: 0 high: 1) weight = 0.12	Weighted average score (U = usefulness score) U = 0.27P + 0.61 L + 0.12F
P1	Spilling of liquid due to liquid expansion	0.7	0	1	0.31
P2	Non-optimal heating due to a faulty thermostat	1	1	1	1
P3	Scalding of the human hand due to hot steam	0.8	1	0	0.83
P4	Spilling of liquid due to seal and gasket rupture	0.5	0	1	0.26
P5	Non-optimal heating due to reduced contact area	0.6	1	0	0.77
P6	Non-optimal heating due to a faulty switch and damaged wire	0.9	0	1	0.36

liquid due to seal and gasket rupture) is the least useful. After application of the metric in every related system, this interpretation will be later generalized.

4.3 Assessment of the creative potential of design problems

In this section, we discuss the third research sub-question: How to assess the creativity (creative potential) of a design problem? Following the method proposed by Sarkar & Chakrabarti (2011), we propose that the assessment of the creative potential of design problems should be computed as a product of the novelty and usefulness scores of the design problems (Eqn. 2):

$$\text{Creative Potential}\,(\text{Creativity}) = \text{Novelty} \times \text{Usefulness} \tag{2}$$

A low score indicates low creative potential, and a high score indicates high creative potential. The ranges proposed are [0, 0.33] – low creative potential, (0.33, 0.67) – medium creative potential and [0.67, 1] – high creative potential.

The steps proposed for assessing the creative potential of design problems are as follows:

1. Identify and collect the design problems associated with the system investigated
2. Convert each into a problem SAPPhIRE as described in Section 4.1
3. Compare the old and new problem SAPPhIRE sets with each other for their novelty scores
4. Create a problem tree map to identify the sets of problems for usefulness assessment

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

5. Assess them for their usefulness scores using the proposed method as described in Section 4.2
6. Compute the creative potential using the creative potential formula provided in Eqn. (2) (Figure 17)

4.3.1 Results

Computation based on the existing case study is given in Table 8. Novelty and usefulness scores are taken from Sections 4.1.1 and 4.2.1, respectively.

According to this calculation, Problem 2 has the highest creative potential (having both high novelty and high usefulness), followed by Problem 5 (although a novel problem, it is less useful to solve). The problem with the least creative potential is Problem 1 and 4, which have a medium novelty but a very low usefulness score.

Correlation between novelty and creative potential is found to be 0.6, between usefulness and creative potential is found to be 0.64 and between novelty and usefulness is found to be 0.02, which is well justified in the literature. This gives further credence to the correctness of the proposed method.

4.3.2 Impact evaluation of the proposed method

To evaluate the proposed method, we plan to design an experiment in which designers will be presented with two sets of design problems within a single context.

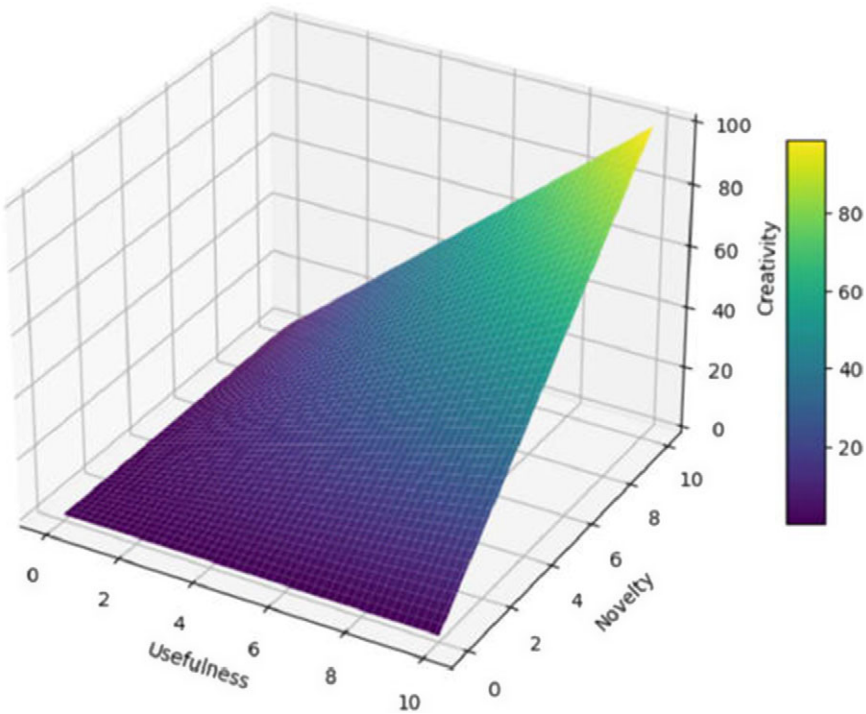


Figure 17. Variation of creativity with novelty and usefulness.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table 8. Calculation of creative potential scores

Problem no.	Problem description	Novelty score	Usefulness score	Creative potential score
P1	Spilling of liquid due to liquid expansion	0.55	0.31	0.17
P2	Non-optimal heating due to a faulty thermostat	0.74	1	0.74
P3	Scalding of the human hand due to hot steam	0.27	0.83	0.22
P4	Spilling of liquid due to seal and gasket rupture	0.65	0.26	0.17
P5	Non-optimal heating due to reduced contact area	0.8	0.77	0.62
P6	Non-optimal heating due to a faulty switch and damaged wire	0.68	0.36	0.25

The first set will comprise problems obtained directly from stakeholders without prior assessment, whereas the second set will consist of problems selected after evaluating their creative potential. Designers will be asked to solve these problems using the brainstorming method, with both the time required to solve each problem (including the time taken for understanding a problem) and the creativity level of the generated ideas at the end of each session being measured. A comparative analysis of the two datasets will then be conducted to examine which set yields greater impact, thereby assessing the overall influence of the proposed method on the design process.

4.4 Discussion

The methods for the assessment of novelty, usefulness and creative potential of design problems are illustrated using a case study of an electric kettle. During the collection of design problems, categories such as safety, reliability and efficiency are considered within the context of the kettle, and these categories are common when identifying design problems across a wide range of engineering and product design domains. The method for evaluating novelty employs the SAPPhIRE model of causality, which is recognized for its general applicability in engineering design, while the variables used to evaluate usefulness have been widely used in the literature across diverse products. By following the proposed steps: collection of design problems, analysis using problem SAPPhIREs (provided the SAPPhIRE representation is feasible) and comparison using semantic similarity for the assessment of novelty, together with the use of three variables in the assessment of usefulness, and consequently the creative potential of design problems, the proposed methods can be useful to a broad range of engineering and product design contexts and are not limited to any specific product type.

5. Summary, conclusions and future work

The overall aim of this research is to develop a method for assessing the creative potential of design problems that can be applied at the early stages of task

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

clarification. This has been carried out in two steps: (1) understand the current state of the art in creativity assessment methods and (2) develop a method for assessing the creative potential of design problems based on novelty and usefulness. In the first step, we reviewed a comprehensive body of literature on creativity and creativity assessment methods, identifying the resources required for method development. In the second step, we worked on three research sub-questions: (1) How to assess the novelty of design problems? (2) How to assess the usefulness of solving design problems? and (3) How to assess the creative potential of design problems?

During the review of literature, we discovered how to identify and analyze design problems by forming the problem SAPPhIRE. The novelty assessment methods proposed by Srinivasan & Chakrabarti (2010b) and Siddharth *et al.* (2019) are the most suitable for the design problems. The adapted research methodology (Singh & Chakrabarti 2025) involved searching for problems in patents, web searches, surveys and questionnaires, converting them to problem SAPPhIREs and comparing them to the existing problem SAPPhIREs for sentence similarity at different abstraction levels. The problem tree was used for the usefulness indicators analysis. Effectiveness evaluation of the proposed novelty assessment revealed that it is more suitable compared to the existing method (Srinivasan & Chakrabarti 2010b). Correctness evaluation revealed relatively high levels of agreement between research findings and experts' intuitive notions.

The proposed method assisted in searching for problems on existing search platforms using keywords derived from SAPPhIRE abstraction levels. The conversion of natural-level descriptions to SAPPhIRE is already a well-established research topic (Bhattacharya *et al.* 2024). The novelty of an idea is determined by the abstraction level at which it exists (Srinivasan & Chakrabarti 2010b); our method required assessing (the various aspects of) a problem for the degree of novelty associated with (each aspect at) each abstraction level. Our method demonstrates that it can be applied to both single- and multi-instance SAPPhIREs (with more than one root cause). The manual process of storing problem descriptions was tedious, so we created a Microsoft Excel tool for efficient storage of problem SAPPhIREs after converting them from natural language descriptions. The time consumption in comparing SAPPhIREs was also high, with each abstraction level of the current problem being compared to all existing problems of the same type. Automating it with a pre-trained database helped reduce the time consumption partially. Expert validation of the proposed novelty assessment method indicated that the proposed method significantly aligns with their intuitive judgment.

The usefulness assessment method is an adaptation of the work of Sarkar & Chakrabarti (2011). AHP is also performed to obtain different weights according to stakeholders' perspectives for the better evaluation of the usefulness of design problems. We calculated usefulness scores for different problems in a variety of contexts by using three of the four parameters they proposed and further quantifying the inputs provided by all stakeholders (designers, customers, etc.). Assessment of usefulness of solving a design problem would help to prioritize those problems that have higher societal value, affect more users and occur more frequently.

Finally, creative potential (Sarkar & Chakrabarti 2011) is assessed. Although the methods are illustrated using a single case study for consistency, the same steps

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

can be applied across a broad range of engineering and product design contexts. This work is intended to help first identify those design problems that are potentially more novel and useful (i.e., more creative) and hence offer greater latitude, subsequently, during problem-solving, in creative exploration of design solutions. Secondly, we argue that the chosen creative design problems among a set of larger problem sets tend to save time and effort by focusing on a problem with the greatest creative potential. We have also proposed a plan for the impact evaluation of the method. Though we plan to conduct it in the future, the correctness evaluation seems to provide some support to the efficacy of the proposed method.

As previously discussed, the proposed method needs to be validated for its impact using the evaluation method and is a part of future work. Our proposed method is currently partially automated; achieving full automation would reduce time consumption and necessitate the development of a web-based tool capable of accepting natural language descriptions of design problems as input and generating complete problem SAPPhIREs for direct comparison with other sets, something the current method does not yet provide and is also a part of our future work.

Acknowledgements

The authors would like to thank all the stakeholders involved in the data collection process, all the anonymous people in the literature for their research work that assisted in curating this research and the experts for their contributions to the validation process.

References

Akin, Ö. 2008 Creativity in design. *Performance Improvement Quarterly* **7**, 9–21.

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

Aru 2025 Artificial intelligence and the internal processes of creativity. *The Journal of Creative Behavior* **59** (2), e1530; doi:https://doi.org/10.1002/jocb.1530.

Basadur, M., **Pringle, P.**, **Speranzini, G.** & **Bacot, M.** 2000 Collaborative problem solving through creativity in problem definition: expanding the pie. *Creativity and Innovation Management* **9** (1), 54–76.

Bhattacharya, K., **Majumder, A.**, **Bhatt, AN**, **Keshwani, S.**, **BSC, R.**, **Venkataraman, S.**, **and Chakrabarti, A.** 2024 Developing a method for creating structured representations of working of systems from natural language descriptions using the SAPPhIRE model of causality. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **38**, e24; doi:10.1017/S0890060424000118.

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

Casakin, H. & **Kreitler, S.** (2005), *The nature of creativity in design: factors for assessing individual creativity*. In *Studying Designers International Conference*, Aix-en-Provence.

Casakin, H. & **Wodehouse, A.** 2021 A systematic review of design creativity in the architectural design studio. *Buildings* **11** (1), 31; doi:10.3390/buildings11010031.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Chakrabarti, A. & **Khadilkar, P.** 2003 A measure for assessing product novelty. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design*. Stockholm.

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.

Cohen, J. A. 1960 Coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**, 37–46.

Feldhusen, J. F. & **Goh, B. E.** 1995 Assessing and accessing creativity: an integrative review of theory, research, and development. *Creativity Research Journal* **8** (3), 231–247; doi: 10.1207/s15326934crj0803_3.

Goel, V. & **Pirolli, P.** 1992 The structure of design problem spaces. *Cognitive Science* **16** (3), 395–429. doi:10.1016/0364-0213(92)90038-V.

Grace, K., **Maher, M. L.**, **Fisher, D.** & **Brady, K.** 2015 Data-intensive evaluation of design creativity using novelty, value, and surprise. *International Journal of Design Creativity and Innovation* **3** (3–4), 125–147; doi:10.1080/21650349.2014.943295.

Guttentag, D. 2013 Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism* **18** (12), 1192–1217; doi: 10.1080/13683500.2013.827159.

Hanna, R. 2018 Computation as 'intervention': a systematic review of 'outcome'. *International Journal of Design Sciences and Technology* **23** (1), 79–108

Hart, S. L. & **Christensen, C. M.** 2002 The great leap: driving innovation from the base of the pyramid. *MIT Sloan Management Review* **44**, 51.

Hendler, J. 2006 Computers play chess; humans play go. *IEEE Intelligent Systems* **21** (4), 2–3; doi:10.1109/MIS.2006.68.

Jiang, F., **Ma, J.**, **Webster, C. J.**, **Chiaradia, A. J.**, **Zhou, Y.**, **Zhao, Z.** & **Zhang, X.** 2023 Generative urban design: a systematic review on problem formulation, design generation, and decision-making. *Progress in Planning* **180**(1–35):100795; doi: 10.1016/j.progress.2023.100795.

Kao, J. J. 1996 *Jamming: The Art and Discipline of Business Creativity*. Harper Business.

Lockhart, E. N. S. 2025 Creativity in the age of AI: the human condition and the limits of machine generation. *Journal of Cultural Cognitive Science* **9**, 83–88; doi:10.1007/s41809-024-00158-2.

Lozano, D. J. 2009 *Metodología para la eco-innovación en el diseño paradesensamblado de productos Industrials (Doctoral dissertation, Universitat Jaume I).*

Maher, M. L. & **Fisher, D. H.** 2012 Using AI to evaluate creative designs. In *DS 73–1 Proceedings of the 2nd International Conference on Design Creativity*, Vol. **1**. The Design Society. ISBN:978-1-904670-39-1.

Miller, K. 2019 Market research. https://medium.com/@mill7828/market-research-88a66742fb9f.

Moss, J. 1966 *Measuring Creative Abilities in Junior High School Industrial Arts*. Monograph 2. (ERIC No. ED022023). Minnesota Tests of Creative Thinking. American Council on Industrial Arts Teacher Education. https://eric.ed.gov/?id=ED022023

Mumford, R. P. & **Gardner, R.** 1994 *Handbook of Creativity* (ed. R. J. Sternberg). Cambridge University Press.

Nidamarthi, S., **Chakrabarti, A.** & **Bligh, T. P.** 1997 The significance of co-evolving requirements and solutions in the design process. *Proceedings of International Conference in Engineering Design, Tampere* **1**, 227–230.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Obieke, C., **Milisavljevic-Syed, J.** & **Han, J.** 2020 Supporting design problem-exploring with emergent technologies. *Procedia CIRP* **91**, 373–381; doi:10.1016/j.procir.2020.02.189.

Obieke, C. C., **Milisavljevic-Syed, J.**, **Silva, A.** & **Han, J.** 2023 A computational approach to identifying engineering design problems. *ASME Journal of Mechanical Design* **145** (4), 041406; doi:10.1115/1.4056496

Pahl, G. & **Beitz, W.** 1996 *Engineering Design: A Systematic Approach*. Springer-Verlag; doi: 10.1007/978-1-4471-3581-4.

Ranjan, B. S. C., **Siddharth, L.** & **Chakrabarti, A.** 2018 A systematic approach to assessing novelty, requirement satisfaction, and creativity. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **32**, 390–414; doi:10.1017/S0890060418000148.

Roozenburg, N. F. M. & **Eekels, J.** 1995 *Product Design: Fundamental and Methods*. 2nd Edition. John Wiley& Sons Ltd.

Roser, H. M. & **Dartnall, W. J.** 2007 Managing new Technologies in Conceptual Engineering Design, demonstrated on innovative positive displacement device. In *Proceedings of the ASME 2007 International Mechanical Engineering Congress and Exposition. Volume 7: Engineering Education and Professional Development*. Seattle, Washington, USA. November 11–15, 2007, pp. 79–84. ASME; doi:10.1115/IMECE2007-43872

Runco, M. A. & **Chand, I.** 1994 Problem finding, evaluative thinking, and creativity. In *Problem Finding, Problem Solving, and Creativity* (ed. M. A. Runco), pp. 40–76. Ablex Publishing.

Sarkar, P. & **Chakrabarti, A.** 2007 Development of a method for assessing design creativity. In *Presented at the International Conference on Engineering Design (ICED'07)*, Paris, France, pp. 79/1–12.

Sarkar, P. & **Chakrabarti, A.** 2011 Assessing design creativity. *Design Studies* **32**, 348–383.

Sarkar, P. & **Chakrabarti, A.** 2015 Creativity: generic definition, tests, factors and methods. *International Journal of Design Science and Technology* **21**, 7–37.

https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

Shah, J. J., **Smith, S. M.** & **Vargas-Hernandez, N.** 2003 Metrics for measuring ideation effectiveness. *Journal of Design Studies* **24** (2), 111–134.

Siddharth, L., & **Sarkar, P.** 2018. A multiple-domain matrix support to capture rationale for engineering design changes. *Journal of Computing and Information Science in Engineering* **18**, 021014-021014-11. https://doi.org/10.1115/1.4039850.

Siddharth, L., **Madhusudanan, N.** & **Chakrabarti, A.** 2019 Toward automatically assessing the novelty of engineering design solutions. *ASME Journal of Computing and Information Science in Engineering* **20** (1), 011001; doi:10.1115/1.4044318.

Siddharth, L., **Chakrabarti, A.** & **Ranganath, R.** 2020 Modeling and structuring design rationale to enable knowledge reuse. *Systems Engineering* **23**, 294–311.

Singh, S. & **Chakrabarti, A.** 2024 Assessment of structuredness of problems in design. *Proceedings of the Design Society* **4**, 1085–1094; doi:10.1017/pds.2024.111.

https://www.spicelogic.com/Products/ahp-software-30

Singh, S., & **Chakrabarti, A.** 2025. Supporting Assessment of Novelty of Design Problems Using Concept of Problem SAPPhIRE. In: Chakrabarti, A., Singh, V., Onkar, P.S., Shahid, M. (eds), *Responsible and Resilient Design for Society*, Volume **4**. ICoRD 2025. Lecture Notes in Mechanical Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-96-5495-6_21.

Srinivasan, V. & **Chakrabarti, A.** 2009 SAPPhIRE—an approach to analysis and synthesis. In *Proceedings of the International Conference on Engineering Design*.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Srinivasan, V. & **Chakrabarti, A.** 2010a An integrated model of designing. *Journal of Computing and Information Science in Engineering* **10**, 031013.

Srinivasan, V. & **Chakrabarti, A.** 2010b Investigating novelty–outcome relationships in engineering design. *Artificial intelligence for Engineering Design, Analysis and Manufacturing* **24**, 161–178.

Stein, M. I. 1953 Creativity and culture. *The Journal of Psychology: Interdisciplinary and Applied* **36**, 311–322; doi:10.1080/00223980.1953.9712897.

Sternberg, R. J. & **Lubart, T. I.** 1999 The concept of creativity: prospects and paradigms. In *Handbook of Creativity* (ed. R. J. Sternberg). Cambridge University Press.

Wang, X., **Nguyen, T.A.**, & **Zeng, Y.** 2015 Influence of information collection strategy in problem formulation on design creativity through mental stress: a theoretical analysis. In: *DS 80-11 Proceedings of the 20th International Conference on Engineering Design (ICED 15)* Vol **11**: Human Behaviour in Design, Design Education; Milan, Italy, 27-30.07. 15. pp. 91–100.

Yin, Y., **Han, J.**, **Huang, S.**, **Zuo, H., and Childs, P.** 2021 A study on student: assessing four creativity assessment methods in product design. *Proceedings of the Design Society* **1**, 263–272; doi:10.1017/pds.2021.27.

Yuan, X. & **Lee, J.** 2013 *Toward a Computational Approach of Creativity Assessment in Product Design*. Computer-Aided Architectural Design Futures.

Yuan, X. & **Lee, J.** 2014 A quantitative approach for assessment of creativity in product design. *Advanced Engineering Informatics* **28**, 528–541.

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Appendix

Table A1. Problem SAPPhIRE database

PS. S. no.	Part	Organ	Input	Physical effect	Physical phenomena	State change	Action
1	Water container	Thermal conductivity of water, specific heat	Less heated element	Convection law	No/less temperature rise of water	Water is not getting heated	
2	Human hand	Skin thermal heating/nonlinear heating	Uncontrolled convection law	Hot steam will start burning human hand	Unburnt to burnt	Burning of hand	
3	Container, switch	Area of top of container	Uncontrolled current	Ampere's law	Uncontrolled current converts into high heat energy	Water-to-steam conversion	Uncontrolled heating/ nonlinear heating
4	Small-sized container	Volume of container	High current	Ampere's law	High current converts into high heat energy leading to quick boiling	Unboiled to boiled liquid	Spilling of liquid
5	Narrow opening area of base of container/ conical shape	Volume of container	Soap/detergent	Principle of surface tension	Soap particles cannot reach the dirt inside the container area	Uncleaned surface remains uncleaned	Less cleaning in base corner
6	Base with nonfilled corners	Area of base	Soap/detergent	Principle of surface tension	Soap particles cannot attack the dirt in the corner area	Uncleaned surface remains uncleaned	Less cleaning in the base corner
7	Corroded heating coil	Element resistance	No/less electric current	Ampere's law	Electrical energy into heat is less	No/less temperature rise of element	Element is not getting heated
8	Corroded steam pipe, faulty bimetallic thermostat	Coefficient of thermal expansion	Less heat from boiling water	Expansion law	Heat energy to expansion is inappropriate leading to less torque to switch	No tripping of switch	Heating will not stop
9	Faulty switch	Switch resistance	No/less electric current	Ampere's law	Electrical energy to heat is less	No/less temperature rise of element	Element is not getting heated

Continued

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table A1. Continued

PS. S. no.	Part	Organ	Input	Physical effect	Physical phenomena	State change	Action
10	Damaged wire	Wire resistance	No/less electric current	Ampere's law	Electrical energy into heat is less	No/less temperature rise of element	Element is not getting heated
11	Kettle base	Base area, base material	Excessive scrubbing	Principle of abrasion of material	Excessive scrubbing leads to abrasion of material	Smooth to rough surface	Localized heating on smooth area
12	Kettle	Coil resistance, volume of kettle	Uncontrolled current	Ampere's law	Uncontrolled current leads to high heat energy generated	Normal to overheated kettle	High energy consumption
13	Narrow opening of container	Area of base, volume of container	Bigger size food items	NA	Bigger size of food items cannot get inserted inside the container	Uncooked remains uncooked	Non-cooking of some food items
14	Bi-metallic thermostat	Coefficient of expansion	Temperature change	Expansion law	Change in temperature increases the strain on thermostats, which changes their dimension	Undeformed to deformed length	Overheating
15	Extension rods	Length, area	Force	Newton's law of motion	Push force from the thermostat will provide longitudinal motion to the extension rod and later to switch	Tripping of switch	Overheating avoidance
16	Kettle body, switch	Thermal conductivity of insulation material	Insulation material	Conduction law	Insulation material over kettle body and switch saves them from melting		Melting avoidance
17	Thermally sensitive materials	Permeability		Principle of thermally sensitive members actuated	Thermal sensitivity actuated due to controlled change in magnetic permeability		Overheating
18	Spout flap, self-closing flaps	Flap insulation	User hand	Ohm's law	Flaps provide insulation to user hands		User safety
19	Cold water chamber	Thermal conductivity	High heat	Conduction law	Heat is getting transferred from the hot chamber to the cold chamber	Cold to hot water	Heating of water
20	Shape memory alloy thermostat	Phase transformation	Temperature change	Law of expansion	Induction of a new phase transformation from martensite to austenite	Hard to soft and ductile	Overheating

Exploring Design Problems

"If you don't see any problem in your surrounding then you are not a good observer"

This questionnaire is intended to collect all the problems which we face directly or indirectly from the equipment we use in our daily lives. This is purely based upon your consent whether you want to fill this questionnaire or not. Kindly do it so that we can use this resource for our research. Thank you for your participation.

While using the below equipment and according to the usage, do you find any problem/s. ? if yes then please state the problem. It can be from within the equipment or can be from your interaction with it. It can be as silly as or as technical as also.

* Indicates required question

Problems???....

1. Name *

Figure A1. Survey questionnaire (for current design problems).

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

2. Age *

3. Occupation *

4. Problems with gas stove? *

5. Problems with gas cutting tools e.g. knife, scissors? *

6. Problems with electric kettles? *

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table A2. Novelty scores as computed from the proposed method for different design problems

S. no.	Problem's description	Novelty scores (least) from proposed assessment	Cumulative decision on novelty
1	Spilling of liquid due to liquid expansion	0.55	Medium novel
2	Spilling of liquid due to faulty seals and gaskets	0.65	Medium novel
3	Spilling of liquid due to the damaged casing of kettle	0.67	Medium novel
4	Non-optimal heating due to damaged wire	0.67	Medium novel
5	Non-optimal heating due to a faulty switch	0.68	Medium novel
6	Non-optimal heating due to reduced contact area	0.8	High novel
7	Scalding of the hand due to hot steam	0.27	Low novel
8	Non-optimal heating due to non-closure of the switch caused by less force from a faulty bimetallic thermostat	0.74	High novel
9	Non-optimal heating due to the uneven, rough surface of the kettle base caused by excessive rubbing	0.63	Medium novel
10	Scalding of the human hand due to hot steam caused by uncontrolled heating	0.6	Medium novel

A1. Assessed design problems

1. Spilling of liquid due to liquid expansion
2. Spilling of liquid due to faulty seals and gaskets
3. Spilling of liquid due to the damaged casing of kettle
4. Non-optimal heating due to damaged wire
5. Non-optimal heating due to a faulty switch
6. Non-optimal heating due to reduced contact area
7. Scalding of the hand due to hot steam
8. Non-optimal heating due to non-closure of the switch caused by less force from a faulty bimetallic thermostat
9. Non-optimal heating due to the uneven, rough surface of the kettle base caused by excessive rubbing
10. Scalding of the human hand due to hot steam caused by uncontrolled heating

A2. Definitions

1. **Purpose (Action)**
The intended goal or function that the system wants to achieve is defined as purpose (in the case of design problems, "is not achieving").

2. **Output (State Change)**
The change of state of a system when it operates is defined as the output of the process (in the case of design problems, "faulty output").

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press

Table A3. Qualitative questionnaire for usefulness score calculation

S. no.	Problem description/variables	Are you encountering this problem?	Do you feel it is important to solve this problem?	How frequently do you encounter this problem?
P1	Spilling of liquid due to liquid expansion	Yes	No	Yes
P2	Non-optimal heating due to a faulty thermostat	Yes	Yes	No
P3	Scalding of the human hand due to hot steam	No	No	Yes
P4	Spilling of liquid due to seal and gasket rupture	Yes	Yes	No
P5	Non-optimal heating due to reduced contact area	No	Yes	Yes
P6	Non-optimal heating due to a faulty switch and damaged wire	No	Yes	No

3. **Process (Physical Phenomena and Effect)**
The underlying natural principles, mechanisms or interactions that enable the system to interact with surroundings are defined as process (in the case of design problems, “problem in phenomena /process”).

4. **Input**
The energy, material or information supplied to the system that activates the process is defined as input (in the case of design problems, “undesirable input”).

5. **Components (Parts and oRgan)**
The physical elements and their structural arrangements that form the system are defined as components (in the case of design problems, “faulty parts/components”).

https://doi.org/10.1017/dsj.2025.10043 Published online by Cambridge University Press