

UNVEILING THE IMPACT OF DESIGN METHODS ON PROBLEM-SOLVING
PERFORMANCE IN UNDERGRADUATE STEM EDUCATIONRoberto Duran-Novoa^{1*} , Felipe Torres-Benoni² ¹Universidad Técnica Federico Santa María (Chile)²Universidad de Santiago de Chile (Chile)*Corresponding author: roberto.duran@usm.cl
felipe.torres.b@usach.cl

Received October 2023

Accepted March 2024

Abstract

Problem-solving is at the core of engineering design, being fundamental for systematic innovation. During their education, students are taught numerous methods and tools, despite that literature shows debatable results regarding their real impact. Consequently, this study aims to quantify the relative impact of design methods on **undergraduate** students' problem-solving performance and determine if this influence goes beyond their background and the problem's complexity.

Utilising novelty, variety, and quality as criteria, the work done by 144 students was evaluated, solving two problems using three methods. The results show a performance improvement of up to 46% when working with methods that guide solution development through design principles. The context and the student method preference did not affect their performance, while the increment in the problem's difficulty improved novelty and variety (15% and 11%) but reduced quality (34%). Surprisingly, the best-performance method was the least preferred, indicating the need of exploring the relationship between performance and actual use.

The study results validate the work invested in teaching design methods, indicating the characteristics of the most efficient ones, beyond expert opinion. The structure of the study allows replication and could help future comparison of results.

Keywords – Design methods, Problem-solving, STEM, Undergraduate education.

To cite this article:

Duran-Novoa, R., & Torres-Benoni, F. (2024). Unveiling the impact of design methods on problem-solving performance in undergraduate STEM education. *Journal of Technology and Science Education*, 14(3), 861-882. <https://doi.org/10.3926/jotse.2473>

1. Introduction

Problem-solving is at the core of engineering design, being fundamental for systematic innovation. During their education, students are taught methods and tools to guide their problem-solving activities, which has shown positive effects on their results (Ananda, Rahmawati & Khairi, 2023; Bourgeois-Bougrine, Buisine, Vandendriessche, Glaveanu & Lubart, 2017; Scott, Leritz & Mumford, 2004). This tendency has also been

supported by less traditional educational approaches, many of them based on active learning principles, which also have its own difficulties and barriers (Valero, 2022).

However, the relative efficiency of the methods taught and their relationship with the specific context or users who participate in each study is still unclear. For example, can it be determined if there is a better method to teach? which criteria could evaluate it? how relevant is the discipline studied? can the conclusions obtained be applied to the workplace?

Most of the studies that compare methods' performance work with a small number of participants and do not follow similar procedures, impeding any statistical analysis, while reaching conflicting conclusions (Cano-Moreno, Arenas-Reina, Sánchez-Martínez & Cabanellas-Becerra, 2021; Chulvi, Royo, Agost, Felip & García-García, 2022; Dumas & Schmidt, 2015; Kannengiesser & Gero, 2015).

Consequently, this study aims to quantify the relative impact of design methods on students' problem-solving performance and establish if this influence goes beyond the student background and problem complexity. The previous could validate the work invested in teaching design methods and point towards which ones to teach.

1.1. Literature Review

Problem-solving is a skill increasingly demanded by society, which requires students to be adequately prepared to engage with today's engineering challenges (Baer, 2015; Bozkurt-Altan & Tan, 2020; Cheng & Yang, 2019; Dumas, Schmidt & Alexander, 2016; Valero, 2022). This has been considered by educational institutions, especially since problem-based learning and its advantages have been studied and applied in multiple contexts (Ananda et al., 2023; Connors, Burns & Campitelli, 2011; Falk, Dym, Agogino, Frey, Leifer, Now et al., 2014; Fiore, Graesser & Greiff, 2018; Ince, 2018; Graesser, Fiore, Greiff & Andrews-Todd, 2018; Priemer, Eilerts, Filler, Pinkwart, Rösken-Winter, Tiemann et al., 2019; Thorndahl & Stentoft, 2020).

Several studies have compared the results differences when solving problems using different methods and parameters, such as creativity effectiveness and idea generation (Genco, Hölttä-Otto & Seepersad, 2012; Linsey, Clauss, Kurtoglu, Murphy, Wood & Markman, 2011a; Scott et al., 2004; Toh & Miller, 2015). Frequently, relevant variables are ignored, like the effect of expert evaluation who assess results using their own tools and judgement (Bourgeois-Bougrine et al., 2017; Chulvi, González-Cruz, Mulet & Aguilar-Zambrano, 2013; Gero, Jiang & Williams, 2013; Mulet, Chulvi, Royo & Galán, 2016). Also, most of the studies cannot be replicated, making it difficult to establish useful areas to explore further and to arrive at general design principles (Crilly, 2015; Daly, Adams & Bodner, 2012), and most of them are carried out with engineering students, despite the possible concentration of projects around similar topics, and the need for multidisciplinary teamwork (Santulli & Langella, 2011).

Another line of studies has sought to determine the impact of design methods more systematically, considering different variables to broaden their analysis (Barbara & Stefano, 2014; Benoni & Novoa, 2023; Chou, 2021). Linsey et al. (2011a,b) assessed the outcomes of brainstorming, brain-sketching, gallery, and C-Sketch / 6-3-5. The problem used for their experiment contained restrictions that increased the challenge (e.g. no electricity available). The output was evaluated considering the quality, quantity, variety, and novelty of ideas (Nelson, Wilson, Rosen & Yen, 2009). Their experiment consisted of 12 teams with six members each. They conclude that methods oriented to group work show better results than those oriented to individual work, pointing out that more experiments would be necessary to determine which methods might yield better performance results in terms of novelty, the variable with the lowest scores. Chulvi (2013) studied the work of designers and engineers attending a doctorate program in design and designers with professional experience (specific field is not indicated). The 48 participants were divided into 16 teams of three people each. Two teams acted as control groups, one using brainstorming, and the other working without method. Half of the 14 remaining teams were assigned SCAMPER, and the other half used the TRIZ Contradiction Matrix. The results were assessed

by a team of experts, who measured novelty and utility as variables, employing the Analytical Hierarchy Process (AHP). Brainstorming, followed by SCAMPER obtained the best results; the no-method team obtained the lowest marks. Hence, employing a method might be better than using none, and among all of them, the most intuitive produces better outcomes. However, they indicate that more investigation is needed, mainly since Brainstorming was used by only one team. Duran-Novoa, Lozoya-Santos, Ramírez-Mendoza, Torres-Benoni and Vargas-Martínez (2019) developed the same experiment in two contexts (countries), adding the possibility of choosing the design method as additional variable. In total, 108 participants were randomly divided into 36 teams of three members each, working with one of four methods: brainstorming, KJ-technique, SCAMPER, and TRIZ. The solutions were measured in terms of novelty, variety, and quality; in both contexts, the most structured methods (SCAMPER and TRIZ) performed better, supporting the idea that methods' impact goes beyond context. It is noteworthy that brainstorming and KJ-technique methods were preferred when participants were allowed to choose how to work, despite producing the most unsatisfactory results. This suggested that, for users, the ease of application and the understanding of the method fundamentals could be more important than generating a better outcome.

In consequence, several considerations could benefit **the** study. First, minimising the effect of evaluators' bias in evaluation is necessary. Second, working in different contexts would lead to identifying the real impact of methods, especially if the results are not consistent. Third, exploring the effect of allowing the participants to select their work method could help academics to choose which design methods to teach, considering which ones are more likely to be applied in the workplace, despite how well they perform when the participants are required to use them under controlled circumstances. Finally, considering the effect of working with problems of different difficulty could help to estimate the extent of methods' benefits in real-world situations.

As far as we know, no study has concurrently considered these variables. There is also a tendency to rely on intuitive approaches and simplify the analysis (e.g. expert opinions, participant interviews) while avoiding statistical analysis, replicability, and potential comparisons. Our intention is to overcome these limitations and deliver reliable results that can be valuable for future studies.

1.2. Research Question

As mentioned at the end of the introduction, it is expected to quantify the impact of design methods on students' problem-solving performance and estimate the extent of this influence. The above can be posed as the following research question: How much impact do design methods have on students' performance when solving problems of different difficulty?

To estimate the extent of this impact, the potential effects of the learning context on students will be considered (the method should surpass it), and the impact of allowing the students to select the method to work with (a preferred method might deliver relatively better results).

2. Methodology

The experiment compared the performance of students when using design methods to solve problems. The participants were divided into four groups, three worked using a design method, and one worked intuitively (see section 2.2). They had to find solutions in a short amount of time, aiming to favour the methods that deliver feasible solutions in a limited time frame (relatively more efficient) and to reduce the effect of non-controlled factors that occur in long experiments, such as influences among groups (Atilola, Tomko & Linsey, 2016), results of other sources of information applicable to the problem (Petre, 2004), and lack of repeatability (Linsey et al., 2011a). For each problem, the methods were assigned to some teams and chosen by others to identify possible relations between participants' preferences and results. The control group did not change throughout the experiment.

2.1. Selected Design Methods

Design methods can be defined as “patterns of behaviour employed in inventing things of value that do not yet exist” (Gregory, 1966); applying them should produce “new design ideas based on the designer’s previous knowledge” (Sakae, Kato, Sato & Matsuoka, 2016).

The chosen methods have been used and studied in previous research, demonstrating a significant impact on user performance. They have been analysed and discussed in design-related conferences and journals and are frequently taught in STEM careers. As a result, although atomized, there is abundant information available to orient our analysis.

2.1.1. TRIZ – The Contradiction Matrix

The Theory of Inventive Problem Solving (TRIZ) proposes that there is a set of universal principles behind all inventions and that these principles can be identified and codified to make the inventive process more predictable (Altshuller, 1984; Belski, 2009; Cavallucci & Oget, 2013). It begins with the abstraction of a specific problem into a typical one (a contradiction), allowing to find a standard TRIZ solution, which should be developed into a specific solution for the particular problem, as presented in Figure 1.

TRIZ has been studied in Chulvi et al. (2013) and Duran-Novoa et al. (2019), working with its most popular tool, the Contradiction Matrix (MaTriz). This selection has the potential to allow comparing overall results and conclusions.

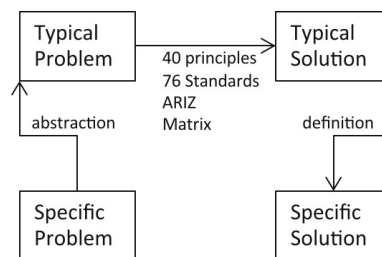


Figure 1. Basic representation of TRIZ's strategy

2.1.2. SCAMPER

SCAMPER is a mnemonic that stands for Substitute, Combine, Adapt, Modify, But to other use, Eliminate, and Reverse. It was proposed by Alex Osborn as the Scamper method and later was denominated the SCAMPER Technique (Eberle, 1996). It establishes an algorithmic structure to solve problems; the user analyses the problem or its parts in simple terms, considering the seven SCAMPER principles to delimit the search areas to stimulate creative ideas for improvement or new developments (Table 1).

Letter	Meaning	General questions
S	Substitute	What can be substituted? Which elements could be added? Sounds? Lights?
C	Combine	What can be blended, mixed, or included? Which parts can be repeated, duplicated, triplicated, etc.?
A	Adapt	What is similar to the “objective?” Does its shape suggest something? A new colour, smell, texture?
M	Modify (also, Magnify or Minimize)	What can be made larger or smaller? Longer or thicker? More compact or shorter?
P	Put to other use	What are the other ways to use it? Does its shape suggest other uses?
E	Eliminate (also Elaborate)	Which parts can be eliminated? What if there is no product?
R	Reverse (also Rearrange)	What happens if the assembly is reversed? Can it be turned inside out? Upside down?

Table 1. SCAMPER questions, adapted from (Gero et al., 2013)

2.1.3. The KJ-Technique

The KJ-technique is a consensus-building method that helps to externalise and prioritise large quantities of ideas and information. It proposes solving problems first through a focus question, followed by creative work conducted through an individual idea-generation process (see Table 2); subsequently, the ideas are collectively analysed, deepening the most promising ones (Kawakita, 1991). Researchers have concluded that this method frequently achieves more and better-quality ideas because it prevents group work factors from reducing the idea-generation potential (Aslani, Naaranoja & Kekale, 2012; Diehl & Stroebe, 1987; Dunnette, Campbell & Jaastad, 1963; Sakae et al., 2016).

The KJ-technique was selected because it has a straightforward procedure that guides the designers' actions. Still, it does not indicate search spaces or solution principles like SCAMPER or the MaTriz. Furthermore, it relates to the previously studied techniques (Duran-Novoa et al., 2019; Kunifuji, 2016; Linsey et al., 2011b; Sakae et al., 2016).

Step	Activity	Description
1	Question	On a board, the team defines the focus question (problem).
2	Alternatives	In silence, members propose alternatives using notes.
3	Grouping	In silence, the team groups similar or related alternatives.
4	Filtering	Duplicated ideas are deleted, while the related ones are linked (e.g., arrows).
5	Development	Each group of ideas is analysed as a potential solution, being expanded, restricted, or decomposed. If necessary, the cycle repeats.

Table 2. The Five Steps of the KJ-technique, adapted from (Kawakita, 1991)

2.2. Participants

In total, 144 undergraduate students aged 19 to 23 participated in this study voluntarily, without compromising their academic work and having the possibility to withdraw at any moment. They were randomly paired into 72 teams of two persons, and later divided into four groups determined by their working method: the MaTriz, SCAMPER, KJ-technique, and the Control group which worked intuitively without method (see table 3). Participants were students from three institutions: Universidad del Desarrollo (UDD), Universidad de Santiago de Chile (USACH), and Universidad de Chile (UCH), Both UDD and UCH were 1st-year students, while USACH students were last year students (6-year program). None of them had experience working with the selected methods, which was validated through a brief survey and interview before the experiment.

After enrolment, each participant was assigned a code to determine the groups and teams, preventing any subsequent identification. The Control group was integrated by UDD students to allow comparing UDD-method teams against UDD-control teams, maintaining unaltered the other variables. The number of participants impeded doing something similar with the other institutions.

Context	Program	Year	Problem-one				Problem-two			
			KJ Tech.	SCAMP.	MaTriz	Control group	KJ Tech.	SCAMP.	MaTriz	Control group
UCH	Architecture	1	6	6	6	0	8	8	2	0
UDD	Engineering, Architecture and Nutrition	1	4	11	9	18	6	12	6	18
USACH	Engineering	6	1	11	0	0	4	4	4	0
Teams by group (methods)			11	28	15	18	18	24	12	18
Teams total			72				72			

Table 3. Distribution of methods and teams (both problems)

2.3. Problems

Both problems were adapted from literature, having enough differences in their proposed challenge to consider the second problem more difficult. Considering the STEM fields (science, technology, engineering, and mathematics), both problems avoid the need of specific technological or mathematical knowledge, as this can create barriers that hinder the understanding of the underlying methods and their impact.

2.3.1. Problem-One: Cutting Board.

Problem-one was “Design a cutting board that can be taken to the dining-room table”, previously employed in Duran-Novoa et al. (2019). The design should consider the kitchen and dining-room requirements (e.g., having mechanical resistance for kitchen use, allowing safe displacement, avoiding liquid spillage in the dining-room, etc.). Problem-one did not have additional restrictions other than serving its function, like cost, specific materials, weight, etc. This was done to create a favourable scenario for the free generation of ideas, reducing the probability of a team not proposing a solution. A scientific approach should yield benefits without being overtly influenced by technological or engineering knowledge.

2.3.2. Problem-Two: Peanut-Shelling Device

Problem-two was “Design a peanut-shelling device, targeting the throughput, separating the peanut shell from the nut, and using non-electrical energy sources”, previously employed in Linsey et al. (2011a,b). Problem-two was considered relatively more complex than Problem-one as it required finding a solution that met the objective within the constraints of energy sources, making it more applicable in a professional context. As a result, having an engineering background or “hands-on” experience was expected to have a positive impact.

2.4. Sample and Procedure

The experiment took place in the Institutions’ classrooms of each context. After receiving a brief introduction to the experiment, the control group was separated while the other participants were trained in the three methods (MaTriz, SCAMPER, and KJ-technique).

After the training, 30 teams were assigned a method to work with before explaining Problem-one, and 24 teams had the option to choose. The same procedure was developed before Problem-two, reversing the proportion (Table 4).

Method	Number of teams (Problem-one)	Number of teams (Problem-two)
Assigned	30	24
Selected	24	30
Control Group	18	18

Table 4. Team distribution by method assigned or selected (both problems)

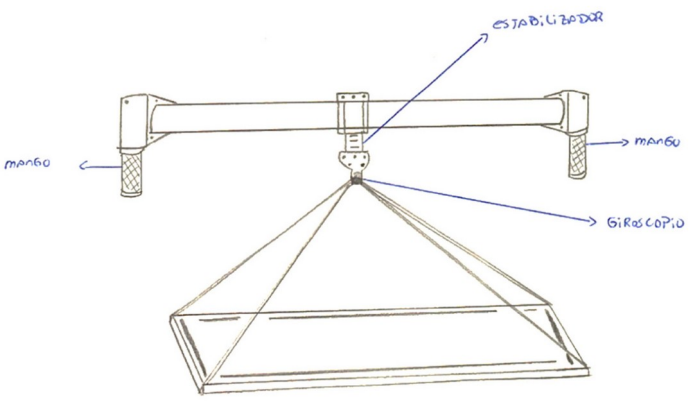
Idea sheet		
Problem: 1	Group: S ₃ (e)	Team: T ₁
Proposal:	Career: PCIC	Semester: 2
<p><u>PRINCIPIO 2B: REEMPLAZO DE SISTEMAS MECANICOS</u></p> 		
<p>Notes: - TABLA DE CORTAR CON ESTABILIZADOR DE MOVIMIENTO CONSISTE EN UNA TABLA ACOPADA A UN [E.D.M] QUE EVITA CUALQUIER TIPO DE MOVIMIENTO, EVITANDO DERRAMES POSIBLES ACCIDENTES DE DERRAME DE LIQUIDO</p>		

Figure 2. Example of a proposal in the Idea sheet; it was obtained working with the MaTriz, specifically the inventive principle “replacement of mechanical systems”

The teams working with methods could choose once for the first or second problem. When the groups had the choosing option, there were no restrictions on the choice; consequently, those teams that had the opportunity to choose in Problem-two could repeat the method used in Problem-one.

Figure 3 shows the general structure of the experiment. In more detail, it was divided into the following six stages:

Stage 1: A brief presentation of the experiment was given to the participants. The materials to be used were provided (e.g., drawing sheets, post-its, blackboards, etc.), and the participants were given general instructions. It was pointed out that we were evaluating the methods’ impact; consequently, rigour and honesty were expected. Groups (three methods, one control group) and participants teams were formed randomly.

Stage 2: Teams were sent to different classrooms with similar conditions. The classrooms were large enough to allow working without influencing another team. Within teams, they decided how to distribute their work.

Stage 3: The teams working with methods were trained in the MaTriz, SCAMPER, and the KJ-technique for 65 minutes. This training used a sample problem to show each method’s practical application, illustrating solution ideas with drawings or prototypes. The problem was “Design an office desk to alternate standing and sitting work”, already studied in Chulvi et al. (2013).

Stage 4: The Problem-one was introduced. The time allowed for participants was 45 minutes. The results were registered in idea sheets, where students drew solution proposals, including any explanation

necessary (Figure 2). Each team could hand in up to two solutions for the problem, which should benefit the more productive teams since they had the possibility of selecting their proposals.

Stage 5: Students were given a 15-minute break. Afterwards, the same teams gathered to prepare to solve Problem-two. The teams with the option of deciding on a method had a maximum of five minutes to communicate their decision, while the teams who did not have the option were notified of their working method.

Stage 6: Problem-two was introduced, following the same rules described in Stage 4.

During the whole experiment, it was verified that the results were obtained using the assigned or selected methods, avoiding mixes or proposals obtained through inspiration that could be later justified by the method.

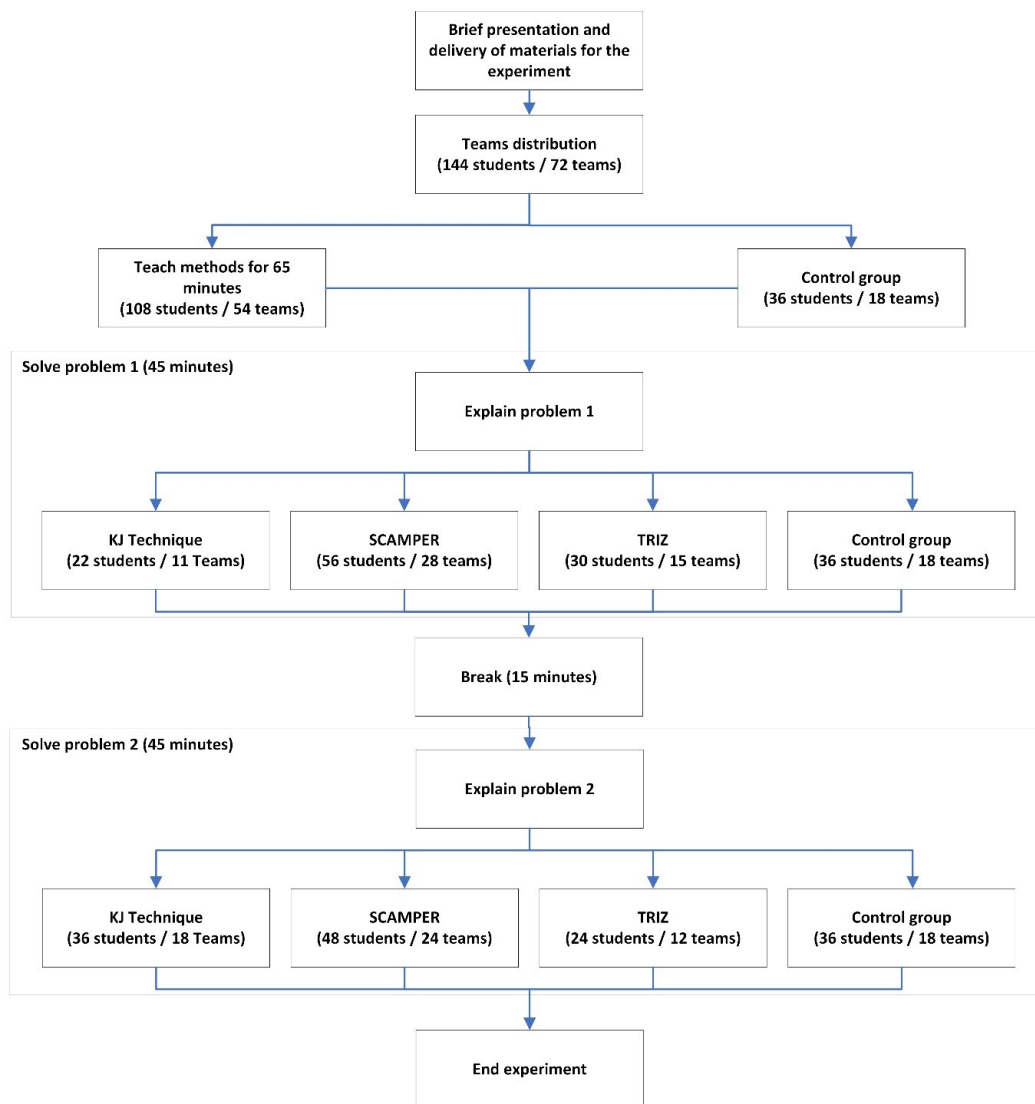


Figure 3. Structure of the experiment

2.5. Evaluation

Teams' proposals were evaluated using a modified version of Nelson's criteria (Nelson et al., 2009), which assigns a holistic score to each proposal based on partial evaluations of *Novelty*, *Variety*, *Quality*, and *Quantity*. The study evaluation omits *Quantity* due to the limited time frame, measuring it indirectly through restricting the number of proposed solutions: best methods should aid their users to be more productive

and deliver more valid proposals to select from (Table 5). Each criterion was assessed on a 5-point scale, ranging from 0 to 4; 12 points was the maximum possible per proposal, 24 per team. Two independent evaluators conducted the evaluations, who had to agree on every score.

Although bias is always possible, it was minimised through the relative evaluation of the proposals. The evaluators scored each solution proposal relative to the 144 proposals received, not based on their prior knowledge. For example, a maximum score in *novelty* can be obtained only if no other team proposes a similar alternative; if the evaluator knows that a similar idea already exists in the market, it does not influence the proposal score. Similarly, if a team proposes an alternative based on electromagnetic properties and another based on fluids, its *variety* should be the highest. Regarding *quality*, all proposals could obtain a four if they are “feasible and easy to implement”, despite the possibility that one proposal could be more accessible to implement than another. Figure 2 presents a proposal that obtained a 4 in novelty (unique), a 4 in variety (different physical principles between proposals), and a 2 in quality (difficult to implement).

Aspect	General description	Measurement	Specific description	Score
Novelty	How unusual or unexpected an idea is	Similarity between the proposed ideas during the experiment	Very common (VC) more than 50%	0
			Common (C) 26%-50%	1
			Uncommon(U) 11%-25%	2
			Very Unusual (VU) 2%-10%	3
			Unique (N)	4
Variety	Explored solution space during the idea generation	Physical principles, working principles (within physics), and proposed embodiments	Same working principles (SWP) and embodiment (E)	0
			Same working principles and different embodiments (DE)	1
			Same physical principles (SPP) and working principles	2
			Same physical principles and different working principles (DWP)	3
			Different physical principles (DPP) and different working principles	4
Quality	Feasibility of an idea and expected performance	Feasibility and ease of Implementation considering cost and utility	Neither feasible (NF) nor easy to implement (NEI)	0
			Hardly feasible (HF) and difficult to implement (HI)	1
			Hardly feasible (HF) or difficult to implement (DI)	2
			Feasible (F) and difficult to implement (DI)	3
			Feasible (F) and easy to implement (EI)	4

Table 5. Evaluation criteria

Each proposal was evaluated following this specific order: first, measuring novelty, then variety, and last, quality, as shown in Figure 4 (the acronyms are described in Table 5). This sequence is more efficient for the evaluators in both scoring and registering. Novelty requires considering all proposals and classifying them to determine their similarities and assign a score; if not evaluated first, the evaluators would have to reclassify ideas after scoring the other criterion. Variety requires both teams' ideas and grouping them facilitates their registration without interfering with the evaluation of quality, which is independent of each proposal.

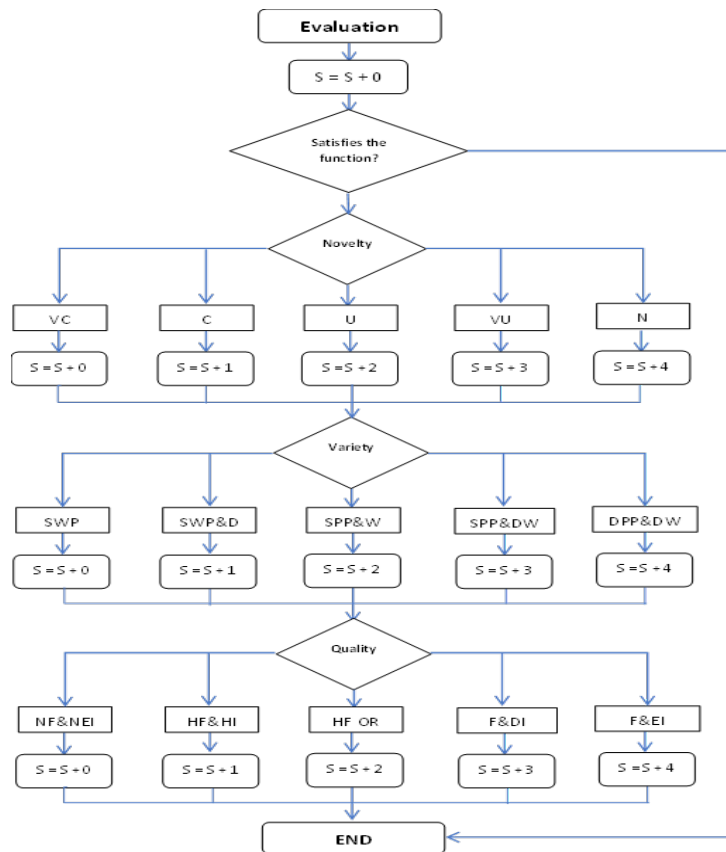


Figure 4. Evaluation process (S means score). For example, the assigned score was immediately zero if the problem function was not accomplished

3. Results

3.1. Problem-One: Cutting Board

There were significant differences in groups performance (Figure 5). The best results were obtained by the MaTriz and SCAMPER, both based on guiding the development of the solution through principles. Figure 5 shows the general score difference when principles are considered (MaTriz and SCAMPER) and when they are not (KJ-technique and Control group); having a similar dispersion, the principle-based methods mean was 54% higher. Table 8 shows examples of students' proposals.

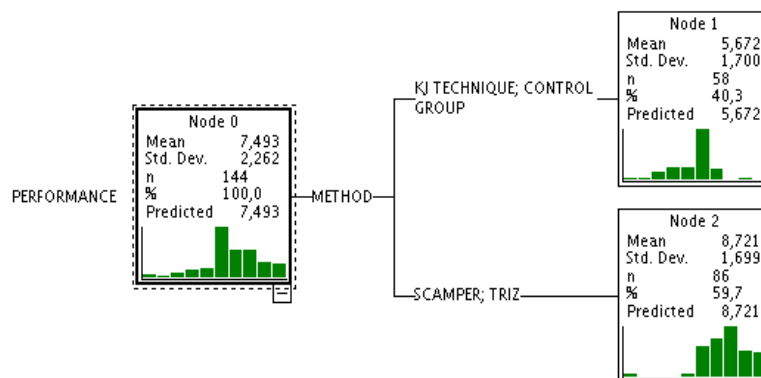


Figure 5. Tree diagram of Problem-one performance, grouping methods as principle-based (MaTriz, SCAMPER) and not (KJ-technique, Control group)

3.1.1. Results by Criteria

Quality obtained the best results and *novelty* the lower. Using four as 100%, their difference is 32% (see Table 6). When methods were chosen, performance declined slightly but not significantly. *Novelty* was also the criterion with the highest dispersion, and where the method's effect was more notorious (Table 7).

Descriptive statistics →	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Novelty	144	0	4	1,85	1,140	1,300
Variety	144	0	4	2,51	0,982	0,965
Quality	144	0	4	3,13	1,026	1,052
Valid N (listwise)	144					

Table 6. Descriptive Statistics (*Novelty*, *Variety*, and *Quality*), Problem-one

		Sum of Squares	df	Mean Square	F
Novelty	Between Groups	61,356	3	20.452	22,983
	Within Groups	124,582	140	0,890	
	Total	185,938	143		
Variety	Between Groups	47,048	3	15,683	24,142
	Within Groups	90,945	140	0,650	
	Total	137,993	143		
Quality	Between Groups	28,855	3	9,618	11,070
	Within Groups	121,638	140	0,869	
	Total	150,493	143		

Table 7. ANOVA table for the criteria, Problem-one

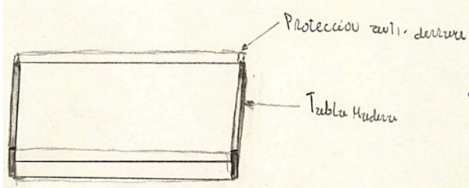
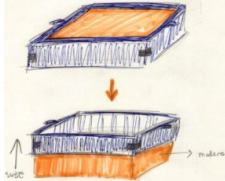
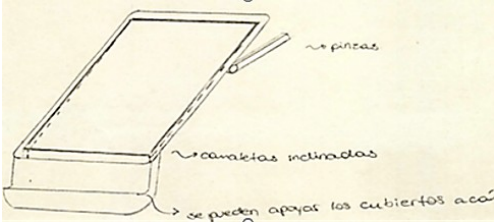
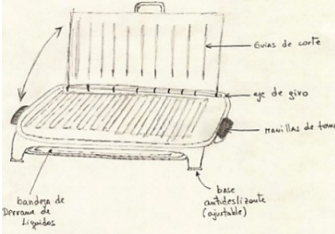
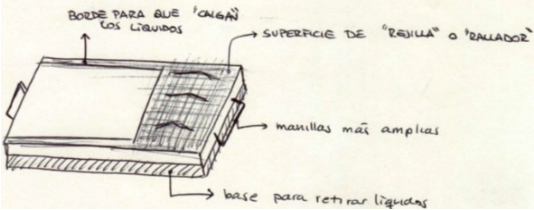
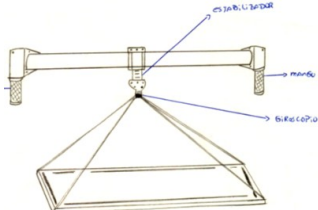
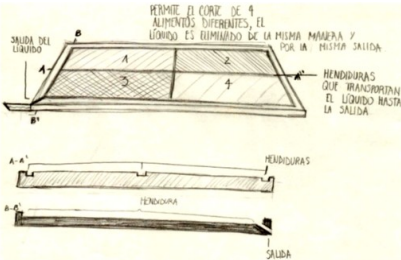
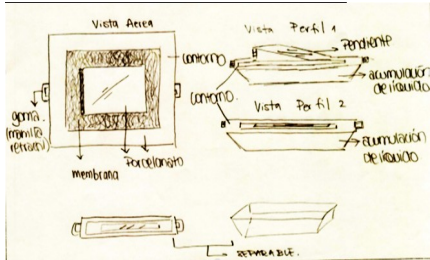
Control Group	KJ-technique
 <p>Anti-spill tray with side container.</p>	 <p>Side barriers with height.</p>
 <p>Tongs with inclined gutters; it contains holder sleeves.</p>	 <p>Cutting Guides, Non-slip Tray with Handles.</p>
SCAMPER	MaTriz
 <p>Liquid retention rim; includes a strainer, a bowl, and handles.</p>	 <p>Inventive principles: Segmentation, taking out, mechanics substitution.</p>
 <p>Flow channels for liquids with reliefs and cutting quadrants.</p>	 <p>Inventive principles: Flexible shells and thin films, colour changes, and parameter changes.</p>

Table 8. Examples of proposals for Problem-one

3.1.2. Results by Methods (Groups)

The best scores were obtained using the MaTriz, across all criteria, followed by SCAMPER and KJ-technique (see Figure 6). The Control group had the lowest scores, independent of group or context.

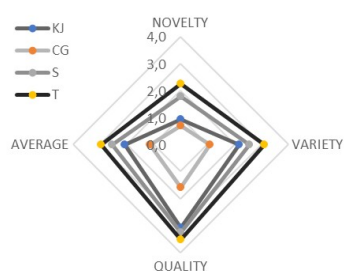


Figure 6. Means for each method, Problem-one

3.1.3. Results by Context

UCH and UDD obtained similar results in all criteria, while USACH -the older students- had a significantly lower *variety* score (similar physics among proposals). Despite having the same background as UDD, the Control group obtained significantly lower scores in all criteria (see Figure 7), a 43% of UDD's average score.

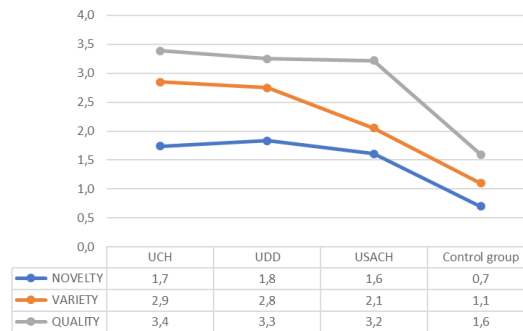


Figure 7. Mean by Context, Problem-one

3.2. Problem-Two: Peanut-Shelling Device

In general, the best results in solving Problem-two were obtained by the MaTriz and SCAMPER, following the same trend observed in Problem-one. In this case, the holistic score difference when applying principles was 44% (Figure 8).

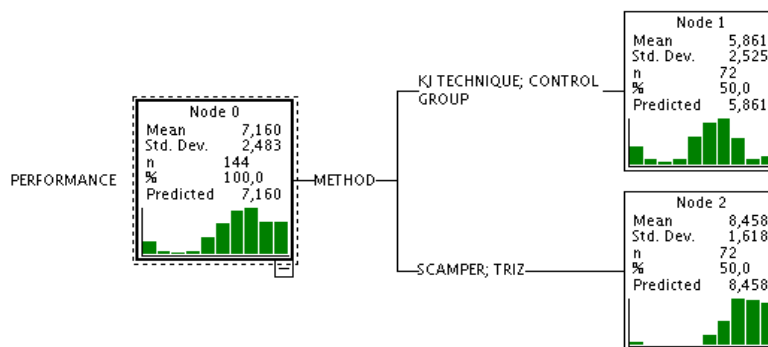


Figure 8. Tree diagram of Problem-two performance, grouping methods as principle-based (MaTriz, SCAMPER) and not (KJ-technique, Control group)

3.2.1. Results by Criteria

On average, *Variety* had higher scores, working with a small number of physical principles but exploring a broader space within it (e.g., working vertically, using fluids, see Table 11). Quality showed the weakest performance (30% less than *Variety*, see Table 9), indicating that proposals were more distant to reality than Problem-one. When methods were chosen, performance was not affected. The method's effect was similar among criteria (Table 10).

Descriptive statistics →	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Novelty	144	0	4	2,46	1,083	1,173
Variety	144	0	4	2,94	1,076	1,158
Quality	144	0	4	1,76	0,948	0,899
Valid N (listwise)	144					

Table 9. Descriptive Statistics (*Novelty*, *Variety*, and *Quality*), Problem-two

		Sum of Squares	df	Mean Square	F	Sig.
Novelty	Between Groups	48,486	3	16,162	18,972	0,000
	Within Groups	119,264	140	0,852		
	Total	167,750	143			
Variety	Between Groups	52,069	3	17,356	21,411	0,000
	Within Groups	113,486	140	0,811		
	Total	165,556	143			
Quality	Between Groups	16,896	3	5,632	7,065	0,000
	Within Groups	111,597	140	0,797		
	Total	128,493	143			

Table 10. ANOVA table for the criteria, Problem-two

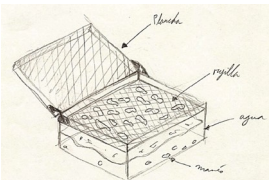
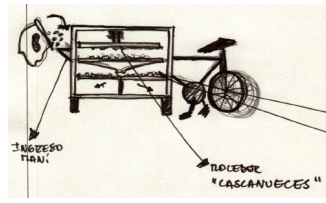
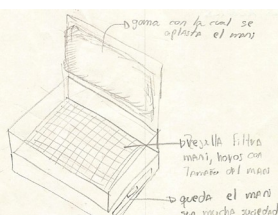
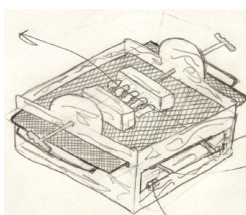
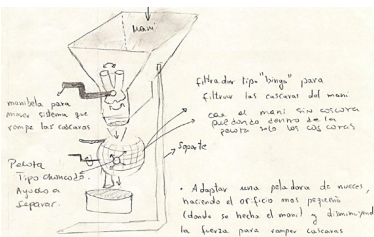
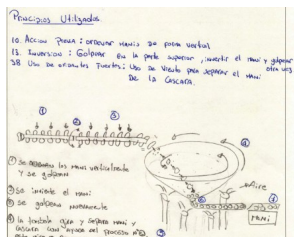
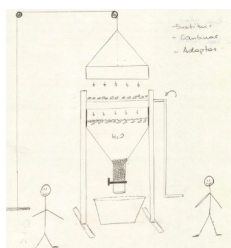
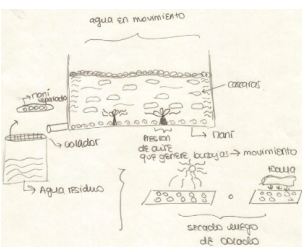
Control Group	KJ-technique
 <p>System: Rack, water, container, and lid</p>	 <p>System: Mechanical energy that drives the peanut to a grinder</p>
 <p>System: Filter with grid and container</p>	 <p>System: Mechanical energy with crusher, sieve, and container</p>
SCAMPER	MaTriz
 <p>System: Silo, handle, holder, separation ball, filter, and container</p>	 <p>Inventive principles: Preliminary action, the other way around, strong oxidants</p>
 <p>System: Silo, water, container, and handle (replace, combine and adapt)</p>	 <p>Inventive principles: Equipotentiality, another dimension, intermediary, mechanics substitution, and thermal expansion</p>

Table 11. Examples of proposals, Problem-two

3.2.2. Results by Methods (Groups)

The MaTriz obtained the best holistic score, followed by SCAMPER, being practically equal in *variety* and *quality* (Figure 9). As in Problem-one, the Control group had the lowest performance in all criteria. Observing the *quality* results of the MaTriz and SCAMPER, it is possible that dealing with more restrictions was facilitated by using methods, but it had the adverse effect of generating unfeasible proposals.

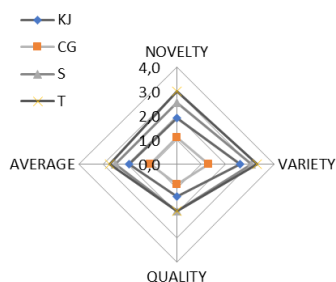


Figure 9. Means for each method, Problem-two

3.2.3. Results by Context

UCH and UDD obtained similar results in all criteria, while USACH performed slightly better, especially in *Novelty* (Figure 10). The Control group had significantly lower scores in all criteria, a 45% of its context reference UDD.

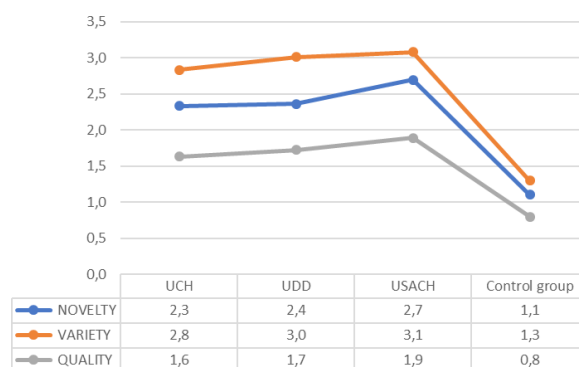


Figure 10. Mean by Context, Problem-two

3.3. Overall Results

The total sample was distributed normally, which changes when the results are observed by each method individually. The MaTriz and SCAMPER showed a bias to higher results (right) that was considered marginal, as the effect of these methods was already known. Thus, the type of student applying the methods and the expert assessing the solutions had no relevant influence (Figure 11).

The best scores were obtained working with the MaTriz. Using its average score as reference (9,23= 100%), SCAMPER reached 89%, KJ-technique 73%, and the Control group 54%.

SCAMPER was the preferred method when given the option. Consequently, the number of SCAMPER teams almost reached the sum of the MaTriz and KJ-technique teams (52 vs 56 teams, see Table 3; 104 vs 112 participants, see Figure 11).

Table 12 presents the average results obtained by each problem considering the criteria. It can be observed that Problem-two had a lower holistic score but not significantly; considering the maximum score as a

reference, less than 3%. However, if the results are analysed by criteria, *quality* has a notorious change, reducing its score a 34%. Novelty and Variety improve their scores, but their sum cannot compensate for the reduction of quality. Figure 12 resumes the relations between problems, methods, and criteria.

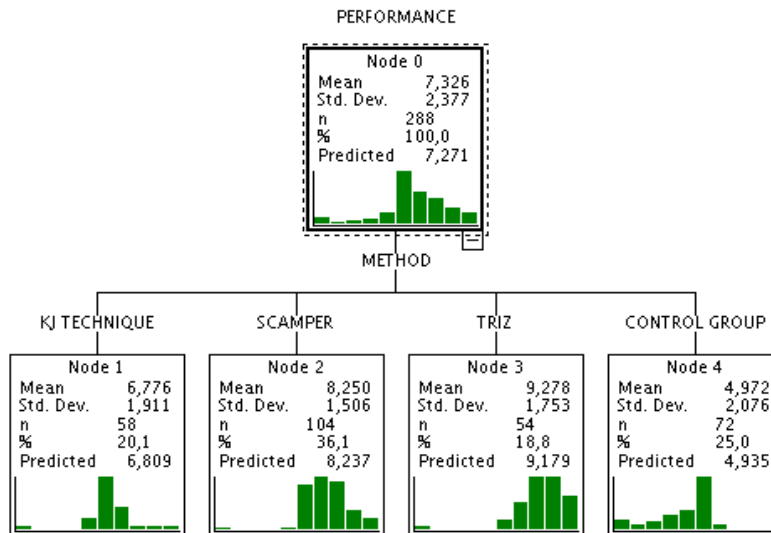


Figure 11. Tree diagram for the methods (both problems)

Criterion	Problem-one	Problem-two	Difference	
Novelty	1,1	1,3	0,61	15,3%
Variety	1,9	2,6	0,43	10,8%
Quality	2,5	3,1	-1,37	-34,3%
Total	7,5	7,2	-0,33	-2,7%

Table 12. Average score-differences between problems by criteria

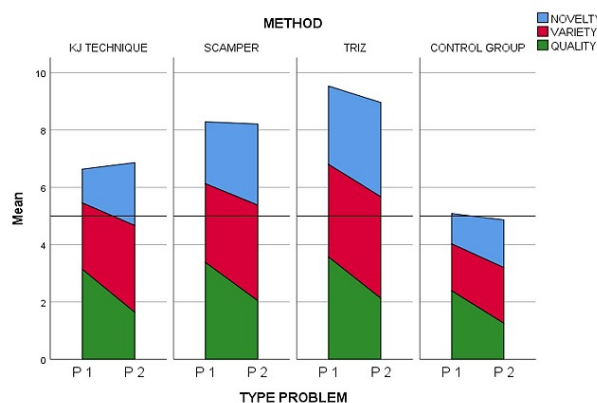


Figure 12. Comparative results (both problems)

Regarding contexts, in Problem-one, UCH obtained better results, while USACH did it in Problem-two. Their different background could have caused this effect, but the differences were not significant.

4. Discussion

The obtained data shows that working with design methods significantly improves the students' performance, independent of the problem difficulty or the student's context. Previous studies reached

consistent conclusions, despite following different methodologies that include direct expert judgement (Gero et al., 2013), working with a small number of participants (Chulvi et al., 2013), working within the same context (Linsey et al., 2011a), or comparing same-difficulty problems (Duran-Novoa et al., 2019).

The methods that utilised principles to suggest or delimit searching areas (SCAMPER and the MaTriz) increased the performance benefit. Equivalent results were obtained by Gero et al. (2013) and Duran-Novoa et al. (2019), where more structured methods produced better outcomes than intuitive ones. Only in Chulvi et al. (2013) a non-structured method -Brainstorming- perform better; however, only one team was using it. It is worth mentioning that SCAMPER was the most preferred method, showing a potentially useful balance between performance and preference.

The MaTriz had the highest scores across problems, contexts, and criteria. The requirement of posing the problem as a contradiction and then studying generic solutions may have the benefit of aiding teamwork. However, despite its superior performance, the MaTriz was the least preferred method when teams had the opportunity to choose. During the experiment, we observed that the teams working with SCAMPER and KJ-technique started proposing ideas quicker than the MaTriz ones, who had to wait until the contradiction stage to discuss proposals; this wait-and-study phase could explain both the better performance and the low preference.

UDD, USACH, and UCH contexts showed similar performances in both problems. The Control group scores were significantly lower, even though its students had the same background as UDD's ones. Both results support the idea that design methods' impact can surpass the learning-context influence and the student's background.

Comparing the obtained results in both problems, the holistic scores did not show relevant differences (less than 3%), but all criteria changed notoriously; *quality* had a score reduction of 34%, while novelty and variety increased theirs (15% and 11% respectively). The probable explanation of this result is that problems of greater difficulty may incentivise or even force innovative and varied ideas but have the undesired effect of reducing their implementation potential (quality). Independent if our explanation is correct, the different results in criteria show the importance of decomposing the analysis; in our case, a holistic approach could conclude that the problem difficulty had no relevance when its effect in *quality* is the strongest one, and any method to be employed should consider it.

USACH students' experience showed a negative effect on problem-one and a positive one on problem-two, none of them relevant. This could happen because when students face basic problems, they may all start on equal footing; however, when problems become more challenging, the students with previous technical knowledge (e.g. manufacturing) can develop quality solutions without affecting their novelty and variety.

An unexpected result was the lack of benefit from choosing the method to work with. Even though some teams repeated the working method in the second problem, no significant differences were observed. This experience acquired by using the method a second time did not deliver observable performance benefits.

4.1. Limitations

Our findings have several limitations to consider, some caused by the specific circumstances of the experiment (unexpected) and others caused by the design of the experiment.

4.1.1. During the Experiment

Initially, teams should provide three proposals per problem, but around a fourth of the teams proposed only two when the experiment was conducted. Therefore, all teams had to select and provide only two proposals, independent of how many they had developed. With more time to develop ideas, some

tendencies could have changed, but this is unlikely since the previously discussed results were observed across all the variables.

Several teams initially avoided working with the assigned method, generating proposals almost immediately. This finding was not quantified, but it did require talking to the students, emphasising the anonymity of their participation and that we were testing the impact of the design methods (something similar was discussed in (Guaman-Quintanilla, Everaert, Chiliza & Valcke, 2022)). Since this problem occurred mainly with the MaTriz, we can speculate that a method's relative inflexibility and difficulty may decrease its quick adoption.

Despite not comparing the effect of drawing or verbalising a solution, drawings were observed to reduce the proposed solution's subjective interpretation, allowing more accurate measurement of the method's performance.

4.1.2. Inherent to the Experiment

Our two-problem approach could point towards relevant considerations when choosing a method, but it is not enough to predict how they could perform in real-life problems. When evaluating the effect of a specific method on a real-life situation, most literature utilises a single case study, which always leaves doubt about the real impact of the method compared with the effect of the circumstances, especially the user. Thus, it is necessary -for example- to develop a standard protocol that can be used during case studies, allowing a meta-analysis.

When measuring the impact of design methods, the learning context was constituted by multiple variables (institution, career, previous knowledge, city culture, etc.). While the data indicates that methods are more significant than the context, there is a chance that a particular aspect of the context could have a substantial impact that is obscured by the insignificance of another one, as it happened with *quality* when comparing both problems' results. For example, if the career is significant but the city is not, the overall average may mask the importance of the career. Although this scenario is improbable, it remains a possibility.

The evaluation protocol needs to be improved. We consider that our approach is going in the right direction but having a larger number of evaluators could help objectivity; for example, five evaluators could allow us to utilise statistical tools to determine their degree of agreement (inter-rater reliability) or simply discarding the two most extreme scores and average the other three.

Finally, it is essential to consider the user's purpose when using a method. For example, in an educational context where there is enough time to learn basic concepts and develop them with examples and discussion, the MaTriz (and the theory behind it) should be the first alternative to consider, but if someone needs to generate valid alternatives in a short amount of time, working with SCAMPER could be a better decision, despite its potential lower performance. Studying the affinity between method and circumstances beyond a case study could provide valuable information to academics and professionals when deciding how to confront their specific challenges.

5. Conclusions

This article aimed to determine how much design methods impact students' performance when solving problems of different difficulties. Our results show that their use significantly improves students' performance, independent of the problem's difficulty or the student's learning context.

The observed benefits increase when the utilised method offers an exploring structure based on design principles and are not affected if the method was assigned or chosen.

A relevant observation is that the method that obtained the best results was also the least preferred. This could indicate the need to explore how performance relates to the actual use of a design method; in our study, performance and preference were not correlated.

Our results and methodology can be used as a reference to compare and determine the real impact of design methods. We hope to see related research comparing findings and results beyond case studies.

Declarations of Conflicting Interests

All authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

- Altshuller, G.S. (1984). *Creativity As an Exact Science* (1st ed.). Gordon and Breach Science Publishers.
<https://doi.org/10.1201/9781466593442>
- Ananda, L.R., Rahmawati, Y., & Khairi, F. (2023). Critical thinking skills of chemistry students by integrating design thinking with steam-pjbl. *Journal of Technology and Science Education*, 13(1), 352-367.
<https://doi.org/10.3926/jotse.1938>
- Aslani, A., Naaranoja, M., & Kekale, T. (2012). Application of creativity techniques in the creation of organizational strategies. *International Journal of Engineering Business Management*, 4, 14.
<https://doi.org/10.5772/51359>
- Atilola, O., Tomko, M., & Linsey, J.S. (2016). The effects of representation on idea generation and design fixation: A study comparing sketches and function trees. *Design Studies*, 42, 110-136.
<https://doi.org/10.1016/j.destud.2015.10.005>
- Baer, J. (2015). The Importance of Domain-Specific Expertise in Creativity. *Roeper Review*, 37(3), 165-178.
<https://doi.org/10.1080/02783193.2015.1047480>
- Barbara, M., & Stefano, F. (2014). Comparison of creativity enhancement and idea generation methods in engineering design training. In Kurosu, M. (Eds.), *Human-Computer Interaction. Theories, Methods, and Tools. (HCI) Lecture Notes in Computer Science* (8510, 242-250).Springer, Cham.
https://doi.org/10.1007/978-3-319-07233-3_23
- Belski, I. (2009). Teaching thinking and problem solving at university: A course on TRIZ. *Creativity and Innovation Management*, 18(2), 101-108. <https://doi.org/10.1111/j.1467-8691.2009.00518.x>
- Benoni, F.T., & Novoa, R.D. (2023). Enhancing Business Innovation: A Review of the Impact of Design Methods. *Journal of Technology Management & Innovation*, 18(4), 104-123.
<https://doi.org/10.4067/S0718-27242023000400104>
- Bourgeois-Bougrine, S., Buisine, S., Vandendriessche, C., Glaveanu, V., & Lubart, T. (2017). Engineering students' use of creativity and development tools in conceptual product design: What, when and how? *Thinking Skills and Creativity*, 24, 104-117. <https://doi.org/10.1016/j.tsc.2017.02.016>
- Bozkurt-Altan, E., & Tan, S. (2020). Concepts of creativity in design-based learning in STEM education. *International Journal of Technology and Design Education*, 31(3), 503-529.
<https://doi.org/10.1007/s10798-020-09569-y>
- Cano-Moreno, J.D., Arenas-Reina, J.M., Sánchez-Martínez, F.V., & Cabanellas-Becerra, J.M. (2021). Using TRIZ10 for enhancing creativity in engineering design education. *International Journal of Technology and Design Education*, 1-26. <https://doi.org/10.1007/s10798-021-09704-3>
- Cavallucci, D., & Oget, D. (2013). On the efficiency of teaching TRIZ: Experiences in a French engineering School. *International Journal of Engineering Education*, 29(2), 304-317.

- Cheng, C., & Yang, M. (2019). Creative process engagement and new product performance: The role of new product development speed and leadership encouragement of creativity. *Journal of Business Research*, 99, 215-225. <https://doi.org/10.1016/j.jbusres.2019.02.067>
- Chou, J.R. (2021). A TRIZ-based product-service design approach for developing innovative products. *Computers & Industrial Engineering*, 161. <https://doi.org/10.1016/j.cie.2021.107608>
- Chulvi, V., González-Cruz, M.C., Mulet, E., & Aguilar-Zambrano, J. (2013). Influence of the type of idea-generation method on the creativity of solutions. *Research in Engineering Design*, 24(1), 33-41. <https://doi.org/10.1007/s00163-012-0134-0>
- Chulvi, V., Royo, M., Agost, M.J., Felip, F., & García-García, C. (2022). How the type of methodology used, when working in a natural environment, affects the designer's creativity. *Research in Engineering Design*, 33(3), 231-248. <https://doi.org/10.1007/s00163-022-00387-y>
- Connors, M.H., Burns, B.D., & Campitelli, G. (2011). Expertise in complex decision making: The role of search in chess 70 years after de Groot. *Cognitive Science*, 35(8), 1567-1579. <https://doi.org/10.1111/j.1551-6709.2011.01196.x>
- Crilly, N. (2015). Fixation and creativity in concept development: The attitudes and practices of expert designers. *Design Studies*, 38, 54-91. <https://doi.org/10.1016/j.destud.2015.01.002>
- Daly, S.R., Adams, R.S., & Bodner, G.M. (2012). What Does it Mean to Design? A Qualitative Investigation of Design Professionals' Experiences. *Journal of Engineering Education*, 101(2), 187-219. <https://doi.org/10.1002/j.2168-9830.2012.tb00048.x>
- Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, 53(3), 497-509. <https://doi.org/10.1037/0022-3514.53.3.497>
- Dumas, D., & Schmidt, L. (2015). Relational reasoning as predictor for engineering ideation success using TRIZ. *Journal of Engineering Design*, 26(1-3), 74-88. <https://doi.org/10.1080/09544828.2015.1020287>
- Dumas, D., Schmidt, L.C., & Alexander, P.A. (2016). Predicting creative problem solving in engineering design. *Thinking Skills and Creativity*, 21, 50-66. <https://doi.org/10.1016/j.tsc.2016.05.002>
- Dunnette, M.D., Campbell, J., & Jaastad, K. (1963). The effect of group participation on brainstorming effectiveness for 2 industrial samples. *Journal of Applied Psychology*, 47(1), 30-37. <https://doi.org/10.1037/h0049218>
- Duran-Novoa, R., Lozoya-Santos, J., Ramírez-Mendoza, R., Torres-Benoni, F., & Vargas-Martínez, A. (2019). Influence of the method used in the generation of valid engineering concepts. *International Journal on Interactive Design and Manufacturing*, 13(3), 1073-1088. <https://doi.org/10.1007/s12008-019-00577-4>
- Eberle, B. (1996). *SCAMPER - Games for Imagination Development*. Prufrock Press Inc.
- Falk, I., Dym, C.L., Agogino, A.M., Frey, D.D., Leifer, L.J., Now, C. et al. (2014). Design Thinking: A Unified Framework For Innovation. *Journal of Engineering Education*, 41(3), 1-5. <https://doi.org/10.1002/j.2168-9830.2005.tb00832.x>
- Fiore, S.M., Graesser, A., & Greiff, S. (2018). Collaborative problem-solving education for the twenty-first-century workforce. *Nature Human Behaviour*, 2(6), 367-369. <https://doi.org/10.1038/s41562-018-0363-y>
- Genco, N., Hölttä-Otto, K., & Seepersad, C.C. (2012). An Experimental Investigation of the Innovation Capabilities of Undergraduate Engineering Students. *Journal of Engineering Education*, 101(1), 60-81. <https://doi.org/10.1002/j.2168-9830.2012.tb00041.x>

- Gero, J.S., Jiang, H., & Williams, C.B. (2013). Design cognition differences when using unstructured, partially structured, and structured concept generation creativity techniques. *International Journal of Design Creativity and Innovation*, 1(4), 196-214. <https://doi.org/10.1080/21650349.2013.801760>
- Graesser, A.C., Fiore, S.M., Greiff, S., & Andrews-Todd, J. (2018). Advancing the science of collaborative problem solving. *Sage Journals*, 19(2), 59-92. <https://doi.org/10.1177/1529100618808244>
- Gregory, S.A. (1966). Design and the Design Method. In Gregory, S.A. (Eds), *The Design Method*. Springer. https://doi.org/10.1007/978-1-4899-6331-4_1
- Guaman-Quintanilla, S., Everaert, P., Chiluiza, K., & Valcke, M. (2022). Impact of design thinking in higher education: a multi-actor perspective on problem solving and creativity. *International Journal of Technology and Design Education*, 1-24. <https://doi.org/10.1007/s10798-021-09724-z>
- Ince, E. (2018). An Overview of Problem Solving Studies in Physics Education. *Journal of Education and Learning*, 7(4), 191. <https://doi.org/10.5539/jel.v7n4p191>
- Kannengiesser, U., & Gero, J.S. (2015). Is designing independent of domain? Comparing models of engineering, software and service design. *Research in Engineering Design*, 26, 253-275. <https://doi.org/10.1007/s00163-015-0195-y>
- Kawakita, J. (1991). *The original KJ method*. Tokyo: Kawakita Research Institute.
- Kunifujii, S. (2016). A Japanese problem-solving approach: The KJ Ho method. *Advances in Intelligent Systems and Computing*, 364, 165-170. https://doi.org/10.1007/978-3-319-19090-7_13
- Linsey, J.S., Clauss, E.F., Kurtoglu, T., Murphy, J.T., Wood, K.L., & Markman, A.B. (2011a). An Experimental Study of Group Idea Generation Techniques: Understanding the Roles of Idea Representation and Viewing Methods. *Journal of Mechanical Design*, 133(3), 31008. <https://doi.org/10.1115/1.4003498>
- Linsey, J.S., Clauss, E.F., Kurtoglu, T., Murphy, J.T., Wood, K.L., & Markman, A.B. (2011b). An experimental study of group idea generation techniques: Understanding the roles of idea representation and viewing methods. *Journal of Mechanical Design, Transactions of the ASME*, 133(3), 031008. <https://doi.org/10.1115/1.4003498>
- Mulet, E., Chulvi, V., Royo, M., & Galán, J. (2016). Influence of the dominant thinking style in the degree of novelty of designs in virtual and traditional working environments. *Journal of Engineering Design*, 4828(March), 1-25. <https://doi.org/10.1080/09544828.2016.1155697>
- Nelson, B.A., Wilson, J.O., Rosen, D., & Yen, J. (2009). Refined metrics for measuring ideation effectiveness. *Design Studies*, 30(6), 737-743. <https://doi.org/10.1016/j.destud.2009.07.002>
- Petre, M. (2004). How expert engineering teams use disciplines of innovation. *Design Studies*, 25(5), 477-493. <https://doi.org/10.1016/j.destud.2004.05.003>
- Priemer, B., Eilerts, K., Filler, A., Pinkwart, N., Rösken-Winter, B., Tiemann, R. et al. (2019). A framework to foster problem-solving in STEM and computing education. *Taylor & Francis*, 38(1), 105-130. <https://doi.org/10.1080/02635143.2019.1600490>
- Sakae, Y., Kato, T., Sato, K., & Matsuoka, Y. (2016). Classification of design methods from the viewpoint of design science. *Proceedings of the DESIGN International Design Conference*, 84, 493-502.
- Santulli, C., & Langella, C. (2011). Introducing students to bio-inspiration and biomimetic design: A workshop experience. *International Journal of Technology and Design Education*, 21(4), 471-485. <https://doi.org/10.1007/S10798-010-9132-6>
- Scott, G., Leritz, L.E., & Mumford, M.D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal*, 16(4), 361-388. https://doi.org/10.1207/s15326934crj1604_1

- Thorndahl, K.L., & Stentoft, D. (2020). Thinking Critically About Critical Thinking and Problem-Based Learning in Higher Education: A Scoping Review. *Interdisciplinary Journal of Problem-Based Learning*, 14(1), 1-21. <https://doi.org/10.14434/IJPBL.V14I1.28773>
- Toh, C.A., & Miller, S.R. (2015). How engineering teams select design concepts: A view through the lens of creativity. *Design Studies*, 38, 111-138. <https://doi.org/10.1016/j.destud.2015.03.001>
- Valero, M. (2022). Challenges, difficulties and barriers for engineering higher education. *Journal of Technology and Science Education*, 12(3), 551-566. <https://doi.org/10.3926/jotse.1696>

Published by OmniaScience (www.omniascience.com)

Journal of Technology and Science Education, 2024 (www.jotse.org)



Article's contents are provided on an Attribution-Non Commercial 4.0 Creative commons International License.

Readers are allowed to copy, distribute and communicate article's contents, provided the author's and JOTSE journal's names are included. It must not be used for commercial purposes. To see the complete licence contents, please visit <https://creativecommons.org/licenses/by-nc/4.0/>.