

A Computational Approach to Identifying Engineering Design Problems

Chijioke C. Obieke, first author

Systems Realization Laboratory, Division of Industrial Design, School of Engineering
University of Liverpool,
Brownlow Hill, Liverpool L69 3GH, United Kingdom
chijioke.obieke@liverpool.ac.uk

Jelena Milisavljevic-Syed, second author

Systems Realization Laboratory, Division of Industrial Design, School of Engineering
University of Liverpool,
Brownlow Hill, Liverpool L69 3GH, United Kingdom
j.milisavljevic-syed@liverpool.ac.uk

Sustainable Manufacturing Systems Centre

School of Aerospace, Transportation and Manufacturing (SATM)
Cranfield University, Cranfield, Bedfordshire MK43 0AL, United Kingdom
j.milisavljevicsyed@cranfield.ac.uk
ASME Membership No. 000102058626

Arlindo Silva, third author

Singapore University of Technology and Design
8 Somapah Rd, Singapore 487372
arlindo_silva@sutd.edu.sg
ASME Membership No. 102035333

Ji Han, fourth author¹

INDEX, Business School,
University of Exeter
Tintagel House, 92 Albert Embankment, London SE1 7TY, United Kingdom
j.han2@exeter.ac.uk

¹ Corresponding author.
Ji Han, University of Exeter
j.han2@exeter.ac.uk

39 **ABSTRACT**

40 *Identifying new problems and providing solutions are necessary tasks for design engineers at early-stage*
41 *product design and development. A new problem fosters innovative and inventive solutions. Hence, it is*
42 *expected that engineering design pedagogy and practice should equally focus on Engineering Design*
43 *Problem-Exploring (EDPE) – a process of identifying or coming up with a new problem or need at the early*
44 *stage of design, and Engineering Design Problem-Solving (EDPS) – a process of developing engineering*
45 *design solutions to a given problem. However, studies suggest that EDPE is scarcely practiced or given*
46 *attention to in academia and industry, unlike EDPS. The aim of this paper is to investigate the EDPE process*
47 *for any information relating to its scarce practice in academia and industry. This is to explore how emerging*
48 *technologies could support the process. Natural models and phenomena that explain the EDPE process are*
49 *investigated, including the “rational” and “garbage can” models, and associated challenges identified. A*
50 *computational framework that mimics the natural EDPE process is presented. The framework is based on*
51 *Markovian model and computational technologies, including machine learning. A case study is conducted*
52 *with a sample size of 43 participants drawn worldwide from the engineering design community in academia*
53 *and industry. The case study result shows that the first-of-its-kind computational EDPE framework presented*
54 *in this paper supports both novice and experienced design engineers in EDPE.*

55

56 *Keywords:* Artificial intelligence; Computer-Aided Design; Conceptual Design; Creativity and Concept
57 *Generation; Data-Driven Design.*

58

59 **1 INTRODUCTION**

60 One of the main tasks of a design engineer at early-stage product design and
61 development is to provide an Engineering Design Solution (EDS) to a societal problem
62 using personal knowledge, experience, and background [1, 2]. Another main task is
63 identifying or conceptualising a new Engineering Design Problem (EDP) [3, 4]. The EDP

64 would deliver societal values once it is solved. Many societal problems, such as “created”
65 and “discovered” problems, are elusive and could remain elusive until discovered by a
66 design engineer [5]. A “created” EDP is a problem that remains unknown or would not
67 exist until conceptualized and translated by an engineer to make it an apparent EDP. A
68 “created” EDP does not have a known formulation, method of solution, or solution. A
69 “discovered” EDP is a type of problem that exists, identified by a design engineer, and
70 may or may not have a known formulation, method of solution, or solution [6, 7]. An EDP
71 triggers ideas for inventive or innovative solutions [8]. The value of an EDS depends on
72 the type of EDP it solves. Hence, both coming up with an EDP and providing an EDS are
73 equally important and are standard expectations in engineering design [9, 10].

74 Therefore, focus on the processes leading to an EDP and EDS is expected in
75 engineering design academia and industry. However, this seems not to be the case.
76 Studies suggest that over the years, the focus is mainly on Engineering Design Problem-
77 Solving (EDPS) – the process of producing an EDS [11, 12]. The process of coming up with
78 an EDP or a need at early stages in design [13], referred to as Engineering Design Problem-
79 Exploring (EDPE) in this research, is rarely practiced or discussed in the literature [14-18].
80 This lack of attention on EDPE over the years has increasing consequences, including 1)
81 limiting the comprehensive capability within engineering design, 2) delayed discovery,
82 innovation, and invention, 3) less inventive solutions/products, 4) lack of development of
83 a specific support tool for EDPE, 5) lack of creativity assessment standard and rewards for
84 EDPE, and 6) decline in EDPE skills within the engineering design community [19-22].
85 Research interests in recent times suggest a need for a strong focus on EDPE. The solution-

86 first approach in engineering design is attracting research interests and seems to be an
87 approach that would necessitate computational facilitation in EDPE. In this approach, an
88 EDS is created first in anticipation of a yet-to-be-identified EDP [27]. This is distinct from
89 the widely known problem-first approach, where an EDS is sought for an EDP. The
90 solution-first approach implies a greater need for a process to facilitate new EDPs that
91 match EDS created beforehand. Albeit studies on the scarce attention and practice of
92 EDPE are few, a study on its determinants, consequences, and mitigation lacks. A study
93 that provides empirical evidence in support of reports on the scarce focus on EDPE also
94 lacks.

95 In this paper, the aim is to investigate the EDPE process and practice for
96 information relating to its scarce practice. This would facilitate possible interventions to
97 encourage EDPE practice in academia and industry. Albeit behavioral interventions like
98 formal teaching/training skills and reward systems for EDPE is possible, there are scholarly
99 indications that EDPE involves heavy thinking and is a difficult, challenging, and time-
100 consuming task [23-26]. Hence, special attention in this paper is on EDPE process activities
101 with natural limitations that necessitate computational support. The models and
102 phenomena related to the natural EDPE process are investigated including the “rational”
103 and “garbage can” models, serendipity, and apophenia phenomena. This contributes to
104 coming up with a first-of-its-kind computational approach and tool presented in this
105 paper that mimic and support the natural EDPE process. To mimic the natural EDPE
106 approach, a Markovian model (MM) is used in synergy with computational technologies,
107 including data mining, machine learning (ML), natural language processing (NLP),

108 duplication recognition, and python programming language. A case study is conducted in
109 three parts with a sample size of 43 participants drawn worldwide from the engineering
110 design community in academia and industry. In the first part of the case study, the aim is
111 to obtain empirical evidence on the lack of focus on EDPE within the engineering design
112 community. The aim of the second part of the case study is to test how closely a
113 computationally framed EDP matches a naturally framed one by a human. In the last part
114 of the case study, the value of the computational EDPE support tool – Pro-Explora V1 (Pro-
115 Explora) presented for the first time in this paper is evaluated.

116 Presented in the following section are the natural EDPE process including possible
117 determinants of its scarce practice. Section 3 is on the methodology used in this paper to
118 come up with a data-driven computational EDPE framework. Section 4 is about the case
119 study data collection methods. Qualitative and quantitative results of the case study are
120 presented in Section 5 and discussed in Section 6. The paper is concluded in Section 7.

121

122 **2 LITERATURE REVIEW**

123

124 **2.1 Models and phenomena related to the natural EDPE approach**

125 The EDPE process is characterized by divergent thinking and decision-making for
126 a new EDP. The “rational” model and “garbage can” model relate to the natural EDPE
127 process. The “rational” model is a formal model of science which postulates that careful
128 analyses of previous problems and theories underpin the discovery of a new problem [28].

129 It supports that a new problem is identified progressively or logically based on gaps in
130 previous problems and theories. The “garbage can” model postulates that a new problem
131 emerges stochastically rather than logically [29, 30]. It supports that a new problem
132 comes up from a stochastic synthesis of previous problems that may not be related. A
133 new problem based on the “garbage can” model is considered more creative than that
134 based on the “rational” model [31]. The “garbage can” model relates to connectionism -
135 a cognitive science concept that likens the connections in computer Artificial Neural
136 Networks (ANN) to natural cognitive ability [32]. The computer ANN contains stochastic
137 and complex interconnected nodes that distribute information for ML.

138 The “rational” and “garbage can” models describe the natural process through
139 which new opportunities, ideas, or concepts are produced. Specifically, they are used to
140 describe the process of coming up with research topics or titles. For example, Alter and
141 Dennis [28] states that: “As faculty, we tend to teach our students a formal “rational
142 model” of science in which research activity is driven by a solid understanding of prior
143 work. Under this approach, research topics emerge from a careful analysis of prior
144 research and theory.” Also, project advice to students is to begin the “search for a suitable
145 problem as soon as possible” [33]. In discussing where a new project is found and how a
146 project is identified, selected, developed, and refined, Dennis and Valacich [29] state that
147 the garbage can model is “a more useful model of how research projects are typically
148 developed” where “the key elements of the project are thrown together into a garbage
149 can, mixed together, and out comes the project.” Also, Martin [30] states that an
150 organization looking for a problem could be imagined as a garbage can where, as

151 “members of the organization generate problems and solutions, they dump these into
152 the garbage can” from which a new problem emerges.

153 Serendipity is a cognitive phenomenon related to EDPE or discovering something
154 new and valuable by chance [34, 35]. It is described as one of the mechanisms of
155 innovation [36], and “connotes the profound ability of finding out valuable things
156 different from those who have been exploring by spending a lot of time or for years” [37].
157 Serendipity occurs when observation by a design engineer triggers unexplored
158 possibilities. Usually, this is based on coincidence with the design engineer’s interest,
159 passion, experience, knowledge, cultural background, and so on. It is reported that the
160 “ground effect” in aircraft is a serendipity discovery [38]. Apophenia is another cognitive
161 phenomenon related to the discovery of a new EDP. It is a natural tendency to see or
162 make meaningful, valuable, invisible connections between unrelated or random data
163 [39]. Apophenia is related to the “garbage can” model and ANN. It could lead to an
164 “invention: creating new, previously unimaginable meanings through accident” [40].
165 EDPE is considered challenging, and findings on some determinants are presented next.

166

167 **2.2 Challenges with the natural EDPE**

168 *2.2.1 Memory limitation in the natural EDPE process*

169 Studies show that the average amount of information the short-term memory can retain
170 and process when exposed to a new concept is 7 ± 2 [41, 42]. The number approaches
171 the minimum with an increased number of syllables in a word during processing of a
172 sequence of words [43]. EDPE, as a cognitive activity, inherently involves processing word

173 sequences, as the “rational” and “garbage can” models suggest. Hence, despite the
174 complexity of the brain, its information-processing capacity is limited [44]. The cognitive
175 demands of EDPE could push cognitive limits and present a level of difficulty, confusion,
176 fixation, and demotivation, which could result in abandonment.

177

178 *2.2.2 Cognitive fatigue in the natural EDPE process*

179 Cognitive fatigue is mental exhaustion resulting from tasks requiring deep thinking and
180 could occur within 30 minutes of commencing a cognitive task [45, 46]. An EDPE process,
181 as explained by the “garbage can” model, involves creating unfamiliar concepts through
182 stochastic information combinations, transformations, and/or explorations [47].
183 Cognitive fatigue could be induced during EDPE tasks and manifest as creative burnout,
184 frustration, and/or tiredness leading to withdrawal from the task [48-52].

185

186 *2.2.3 Insufficient knowledge and difficulty in prompt initiation of natural EDPE process*

187 It could be inferred from the “garbage can” model that EDPE requires knowledge and
188 information [53]. Sometimes it could be difficult to promptly recall previous knowledge
189 or information. Hence, to “create” a new EDP by stochastically recalling and combining
190 previous knowledge and information could be challenging. Also, knowledge is infinite and
191 could hinder EDPE irrespective of the level of experience of a design engineer. Hence,
192 novice design engineers who are not experienced, knowledgeable, or informed could find
193 it more challenging to practice EDPE.

194

195 **2.3 Potential computational technologies for EDPE**

196 A computational system is not susceptible to the challenges associated with the
197 natural EDPE process discussed in Section 2.2 [54]. Hence, it could be used to support
198 EDPE. The “garbage can” model discussed in Section 2.1 could be computationally
199 mimicked by using Artificial Intelligence (AI) [55]. A possible approach is to create a data
200 (word) sequencing model using a Markovian model (MM), Bidirectional Encoder
201 Representations from Transformers (BERT), or Long Short-Term Memory (LSTM). An MM
202 is any model that exhibits the Markov chain (MC) property. BERT uses transformers -
203 unique neural network architecture reported effective in modelling long-term
204 dependencies in a sequence [56, 57]. BERT, like LSTM, uses long-term dependencies
205 (depending on previous states) in its network to predict the next state in a sequence. For
206 example, if the state represented with the ellipses is missing from the sequence -
207 “Engineering design is a noble...”, contextual word embeddings BERT can be used to
208 predict (natural language inference, NLI) the next or missing state (masked word) as
209 “profession”. The next state is inferred relative to the previous states. BERT can also
210 combine the context of both previous and next states (bidirectional) to predict a state in
211 a sequence. This could be described as forward and backward determinism.

212 The persistence of previous states in BERT and LSTM networks could be
213 disadvantageous because it can impact computational power, time, speed, and cost. Also,
214 BERT requires significant data training. Although there are indications that BERT can be
215 used to predict/generate a new sequence, it is used to predict/generate a sequence that
216 already exists such as known answers to questions. It is reported that in sequence

217 generation, BERT is formulated as a Markov random field language model without
218 additional parameters or training [58]. When the generation of a new concept is intended
219 such as in EDPE, the repetition of already known concepts or solved EDP is not desirable.

220 Unlike BERT and LSTM, MC is a model of a specific type of stochastic sequence [59,
221 60]. It is memoryless of past states, and the next state in the sequence only depends on
222 the present state. The MC network uses only forward determinism and does not retain
223 previous information. This results in lesser computing power, cost, time, and higher
224 computing speed. In deterministic chaos theory, forward determinism is considered more
225 important than backward determinism [61]. MC has prior stochastic and decision-making
226 applications in engineering design including 1) modelling of transitions in
227 communications, 2) modelling of sequential design decisions, and 3) analyses of
228 behavioral patterns in engineering design [62-65]. An MM is used in this paper as the
229 pioneering technology to provide the basis for comparison with BERT and LSTM in future
230 research in EDPE applications.

231 Big data could be equivalent to the previous knowledge required in the “garbage
232 can” model. Big data is a large volume of structured, semi-structured, and unstructured
233 data from which new knowledge or information can be found [66]. Big data is associated
234 with some computational technologies, including data extraction, NLP, ML, and
235 duplication recognition [67]. Presented in Section 3 is how an MM is used in synthesis
236 with big data and the associated technologies to come up with computational support in
237 EDPE.

238

239 **2.4 Research questions**

240 Despite being an important and standard requirement, EDPE lacks practice within the
241 engineering design community, as discussed in Section 1. There are scholarly opinions
242 that EDPE is a challenging task with cognitive limitations (Section 2.2). Unlike EDPS, there
243 is no specific support tool for EDPE. However, emergent computational technologies
244 could provide support for EDPE. To this end, the research questions (RQs) that follow are
245 addressed in this paper. The answers to the RQs would provide knowledge for
246 interventions in the lack of EDPE practice in engineering design. In the first part of the
247 case study presented in this paper, *RQ1* is addressed. In the second and third parts of the
248 case study, *RQ2* is addressed:

249 ***RQ1:*** Do design engineers understand that EDPE is a standard requirement in
250 engineering design, like EDPS?"

251 ***RQ2:*** Could emergent computational technologies support novice and
252 experienced design engineers in EDPE?

253

254 **3 METHODOLOGY**

255 **3.1 Theoretical framework for computational EDPE**

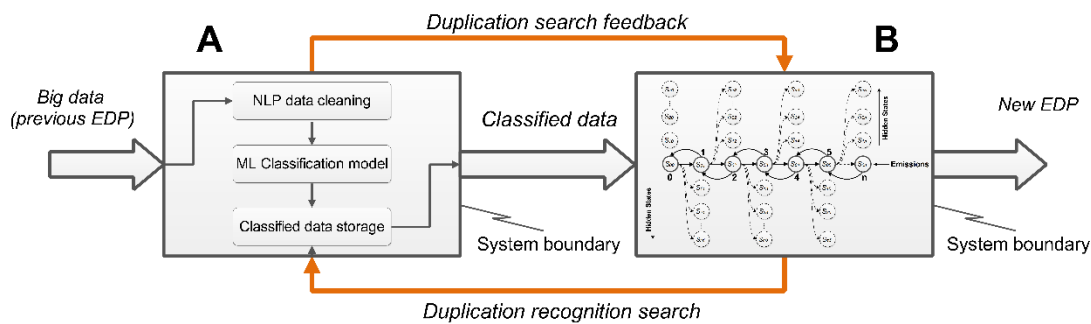
256 Concerning *RQ2*, a theoretical data-driven computational framework, shown in
257 Fig. 1, is presented to support EDPE as a challenging activity. The framework is based on
258 the information presented in Table 1 on the natural EDPE approach and its computational
259 equivalence. The natural EDPE approach in Table 1 is based on findings discussed in
260 Section 2 and supported by most intellectual property offices [68].

261

Table 1 Comparison of algorithms for natural and computational EDPE

Natural EDPE approach	Computational EDPE approach
<p>1) Identify an EDP of societal relevance by accident (serendipity), stochastic synthesis (“garbage can” model), logical progression (“rational” model), and/or conceptualization (apophenia)</p> <p>2) Search manually for prior existence in relevant databases using search engines.</p> <p>3) Decide, subject to acceptance by society or a relevant authority</p>	<p>1) Frame an EDP of societal relevance by stochastic synthesis of big data, computational technologies (data extraction, ML, NLP), coding capabilities, connectionist theory, deterministic chaos, MM, BERT, and/or LSTM</p> <p>2) Make an automated search for prior existence in relevant databases using duplication recognitions.</p> <p>3) Decide, subject to a design engineer's acceptance.</p>

262



263

Fig. 1 Big data-driven computational EDPE framework

265

266 The computational EDPE framework shown in Fig. 1 comprises *System A* and

267 *System B*. The input to *System A* is a collection or corpus of engineering design project

268 titles extracted online from Compendex, Scopus, journals and conferences databases, and

269 findaphd.com using a python data extraction tool “Scrapy”. The output from *System A* is

270 processed data which feeds *System B* to produce a new EDP as output. The project titles

271 used as input to *System A* are important lines of words that represent EDPs in engineering

272 design projects [69]. A “project title should provide information about the topic being

273 studied, and may consist of the actual problem statement” [70]. According to Martin [30],
274 “the researcher should critically review the literature on a given topic in order to find an
275 important issue which previous research has failed to resolve successfully”. This
276 ‘important issue’ is usually formulated as a title – the problem solved, being solved, or yet
277 to be solved. Hence, the extracted titles in this paper are previous EDPs. As an example,
278 in the project title: “Design of an Automatic Sprinkler Fire Fighting System”, the EDP of
279 fighting the fire using an automatic sprinkler system is described. The underlying EDP is:
280 “What is a better way, among the existing alternatives, to fight fire?” The design project
281 is conceived as a solution to fire fighting. It is believed that such a solution was not
282 available when the project was conceived. This makes the project an EDP that needs to
283 be addressed because there is a potential benefit in doing so. The EDP addressed in the
284 project is described with the title. Albeit the title appears as an EDS, it is an EDP if that
285 EDS is unavailable or yet to be realized. It remains an EDP until it is solved. Hence,
286 computationally exploring and identifying a new EDP could lead to an invention or
287 innovation.

288 The increasing volume of titles, continuously collected for EDPE, could be
289 regarded as structured big data of previous EDPs. Hence, this is the concept of the ‘Big
290 data’ indicated as input to the model in Fig. 1. To mimic the natural EDPE, the framework
291 in Fig. 1 is created to use or learn from only the natural EDPs to come up with a
292 computational equivalent. The output from *System B* (which contains the MM in Fig. 2) is
293 a unique EDP distinct from the input. The preprocessing and processing of the corpus in
294 *System A* for input in *System B* to produce a unique EDP are presented next.

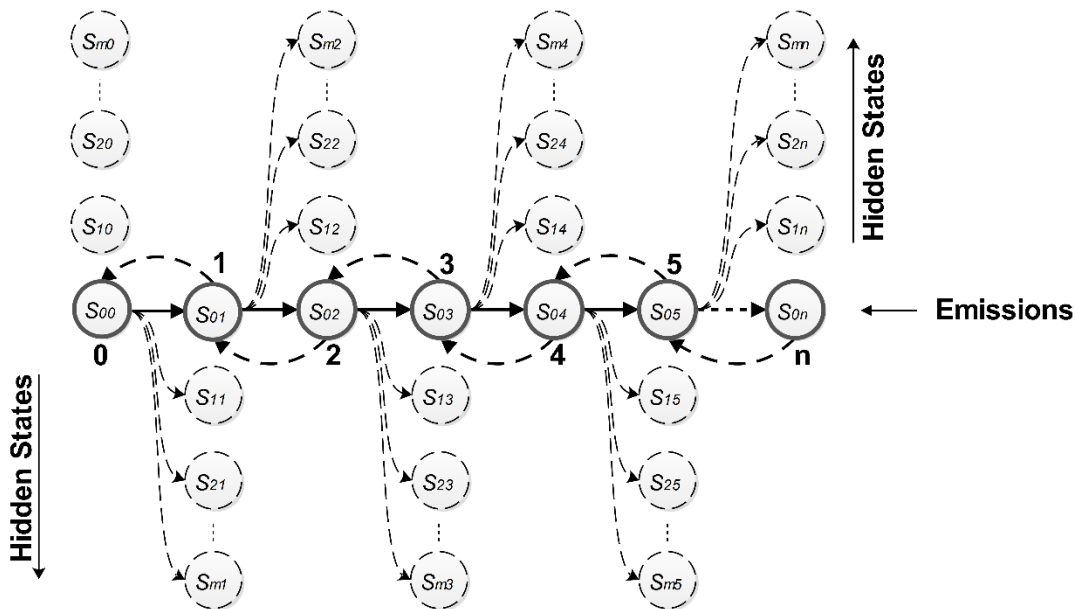
295 3.1.1 *Preprocessing of corpus for EDPE*

296 The corpus extracted online is preprocessed in *System A* Fig. 1 using NLP and ML.
297 The corpus is first prepared as a “tab-separated value” with each line in the corpus ending
298 with a period. On inspection, some of the extracted titles in the corpus appear vague to
299 describe an EDP. This necessitates an ML classification model to classify subsequent
300 extracted titles that do not describe an EDP (Non-EDP). The extracted corpus is manually
301 separated by inspection as a dataset of EDP and non-EDP. This is to enable the training of
302 the algorithm for the classification model using supervised ML – an aspect of AI that
303 provides computer systems with the ability to learn from data. The dataset size is 2133
304 (comprising 1833 EDP and 300 non-EDP), and a 20% test size is used for the ML. The
305 training requires that the dataset is ‘cleaned’ and ‘tokenized’ as part of NLP [71, 72]. The
306 ‘cleaning’ requires the removal of regular expressions or characters that specify a search
307 pattern in extracted texts such as “?”, “@”, and “\$”. It also requires the removal of
308 stopwords from the dataset such as “a”, “for”, and “the” which are insignificant in NLP.
309 Different algorithms are tried during the training, including RandomizedSearchCV, Naïve
310 Bayes (Gaussian and Multinomial), and Random Forest. These algorithms are part of
311 Scikit-learn – a library in Python that provides many unsupervised and supervised learning
312 algorithms. Trying different algorithms to select the best based on performance is a
313 common practice in ML. Two Scikit-learn performance evaluation libraries – classification
314 report and confusion matrix [73], are used during the training to evaluate the
315 performance of the algorithms in the classification model. For RandomizedSearchCV, the
316 accuracy calculated using the confusion matrix is 93%. The classification report shows the

317 precision, recall, and f1-score accuracy metrics as 94%, 93%, and 93%, respectively. These
 318 metrics suggest that only a few EDPs are wrongly classified as non-EDPs and vice versa.
 319 Hence, RandomizedSearchCV is used based on best performance values. The
 320 preprocessed corpus in *System A* Fig. 1 is stored and processed in *System B* Fig. 1 to
 321 produce a new EDP as presented next

322

323 **3.1.2 Processing of corpus for EDPE**



324

325 **Fig. 2 A theoretical model for EDP sequencing [8]**

326

327 The MM contained *System B* Fig. 1, discussed in Section 2.3, is enlarged in Fig. 2
 328 for clarity. In Fig.2, the MM exhibits the properties of a two-stage MC known as a hidden
 329 Markov model (HMM). It has hidden and physically observable states (emissions) [74, 75].
 330 What constitutes the hidden states and emissions in this paper are explained in Section

331 3.3. MC is used as an HMM in many real-life problems, such as handwriting recognition,
 332 machine maintenance, and weather forecasting. This is because MC alone does not fully
 333 represent the intent in many real-life problems [76]. In this paper, as shown in Fig. 2, the
 334 hidden states (for example, $S_{10}S_{20}... S_{m0}$) are discreet, while the emissions $S_{00}S_{01}S_{02}... S_{0n}$
 335 exhibit the Markov property (Section 2.4). The probability of the emissions depends on
 336 the probability of the hidden states. The emissions transit such that the next emission
 337 depends on the present emission and not on the past emission(s). The emission
 338 transitions are assumed to be observed at equal time intervals at the indices $0,1,2,3,... n$,
 339 known as epochs [77]. This means that time-homogeneity (discrete instead of real-valued
 340 time) applies. In this application, the actual time for transitions is computationally very
 341 small and justifies the time-homogeneity assumption. The stochastic process in Fig. 2
 342 could be expressed as Eq. (1).

$$343 \quad f: S_n \times \Lambda \rightarrow S_{n+1} \quad n \in \mathbb{Z}^+; \mathbb{Z}^+ = \{0, 1, 2, 3, \dots\} \quad (1)$$

344 In Eq. (1), as used in this paper, the output S_{n+1} is a function of two arguments, s_n
 345 and λ . This is such that, $s_n \in S_n$ and $\lambda \in \Lambda$. The function $f(s_n, \cdot)$ is a random variable (S_{n+1})
 346 for each $s_n \in S_n$, while for each $\lambda \in \Lambda$, $f(\cdot, \lambda)$ is a hidden function between S_n and S_{n+1} . This
 347 hidden function, $f(\cdot, \lambda)$, makes the output of *System B* Fig. 1 to be indeterministic which
 348 represents the unpredictability of the highest order [78]. Hence, the probability of any
 349 predictor's confidence in the output of *System B* Fig. 1, relative to the input, cannot be
 350 unity. Generally, the transition probability of the MM in Fig. 2 is given in Eq. (2).

$$351 \quad p_{ij} = P(S_{0(n+1)} = T_j | S_{0n} = T_i) \quad (2)$$

352 Eq. (2) shows that, in Fig. 2, if the probability of an emission at epoch n is T_i then there is
 353 a probability that at epoch $n+1$ the emission is T_j [79]. Note that the emission transitions
 354 occur if, and only if, $p_{ij} > 0$. As shown in Fig. 2, an emission is represented as S_{ij} where the
 355 row vector is the sequence index and the column vector is the epoch. For simplicity of
 356 analysis, let the emissions at epochs $0, 1, 2, \dots, n$ in Fig. 2 be a, b, c, \dots, z . Specifically, applying
 357 Eq. (2) to Fig. 2 yields Eq. (3).

$$\begin{aligned}
 & P(S_{01} = b, S_{02} = c, \dots, S_{0n} = z | S_{00} = a) \\
 & = P(S_{00} = a)P(S_{01} = b | S_{00} = a)P(S_{02} = c | S_{01} = b) \dots \\
 & \quad \dots P(S_{0n} = z | S_{0(n-1)} = y) \\
 & = p_a^{(0)} p_{ab} p_{bc} \dots p_{yz}
 \end{aligned} \quad \left. \vphantom{\begin{aligned} & P(S_{01} = b, S_{02} = c, \dots, S_{0n} = z | S_{00} = a) \\ & = P(S_{00} = a)P(S_{01} = b | S_{00} = a)P(S_{02} = c | S_{01} = b) \dots \\ & \quad \dots P(S_{0n} = z | S_{0(n-1)} = y) \\ & = p_a^{(0)} p_{ab} p_{bc} \dots p_{yz} \end{aligned}} \right\} \quad (3)$$

359 Relating Eq. (3) to Fig. 2, $p_a^{(0)}$ represents the initial emission probability for the stochastic
 360 process at epoch 0 . p_{ab} represents the transition probability from epoch 0 to epoch 1 . p_{bc}
 361 represents the transition probability from epoch 1 to epoch 2 . p_{yz} represents the
 362 transition probability from epoch $n - 1$ to epoch n . The initial emission would be necessary
 363 if the observable state of the sequence after n transitions is of interest. If the initial
 364 emission probability is known, then $p_a^{(0)} = 1$, and Eq (3) results in Eq. (4).

$$= p_{ab} p_{bc} \dots p_{yz} \quad (4)$$

366 After an EDP is produced in *System B* Fig. 1 as a sequence of emissions, a
 367 duplication recognition search is performed. The sequence is only outputted from *System*
 368 *B* Fig. 1 as a new EDP if it does not have a duplicate (exact match) in the original corpus
 369 stored in *System A* Fig. 1. Otherwise, it is discarded and another sequence produced. The
 370 framework in Fig. 1 is deployed to produce a computational EDPE tool discussed next.

371

372 **3.3 Pro-Explora – a computational support tool for EDPE**

373 Pro-Explora is a computational support tool for EDPE. In Pro-Explora, the
374 theoretical EDP sequencing model in *System B* Fig. 1 is realized by processing the input
375 corpus in *System B* Fig. 1 as a python dictionary data structure. The corpus is split into
376 single words with each word as the dictionary key. The value list of each key contains all
377 words that come immediately after the key in all occurrences of the key in the corpus. To
378 closely mimic a natural EDP, the initial word/emission at epoch 0 in Fig. 2 is randomly
379 selected from the list of hidden states $S_{10}S_{20}... S_{m0}$. The hidden states $S_{10}S_{20}... S_{m0}$ comprise
380 the first words of each EDP in the extracted corpus. For example, S_{10} will be “Design” and
381 S_{20} will be “A” for a corpus that contains the two EDPs – [“Design of a mechanical intrusive
382 force detection device.”, “A design of an automatic bottle opener.”]. The dictionary key
383 “of” in the corpus will have the values “a” and “an”. After the initial emission, the rest of
384 the emissions are constrained to be randomly chosen from the mutually exclusive hidden
385 states at epochs $1,2,3,4,5,...n$ based on the Markov property. These observable states in
386 Fig. 2 represent the new EDP from *System B* Fig. 1. The framework in Fig. 1 mimics the
387 natural EDPE (Section 2.1), especially the “garbage can” model. As presented next, a
388 useful output from *System B* Fig. 1 requires specific adjustments to the MM in Fig. 2 [76].
389

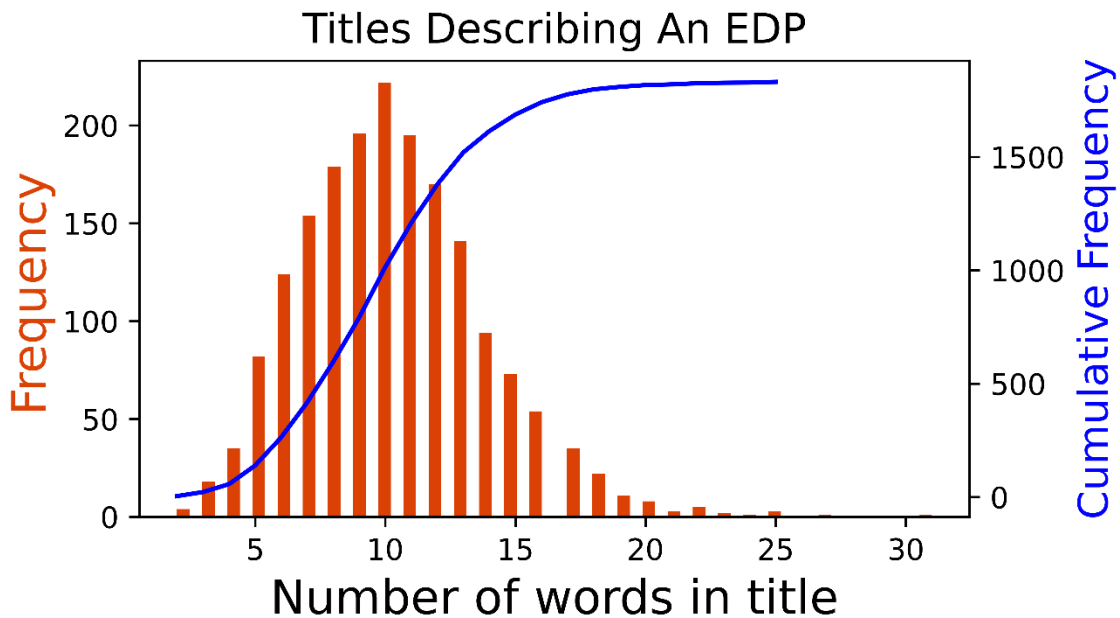
390 **Adjustment 1: Enhanced stochasticity**

391 For a word ending with a period in a list of hidden states in Fig. 2, synonyms of the word
392 are added to the list without replacing the word. For example, “extractor.”, “centrifuge.”,
393 and “threshing machine.” are added to the list containing the word “separator.” The

394 synonyms are obtained using the python wordhoard library and online thesaurus. This
395 adjustment increases the stochasticity and number of Pro-Explora outputs with the same
396 EDPs in the corpus and makes the tool create new EDPs that involve newer technologies.
397

398 **Adjustment 2: Output word-count constraints**

399 The number of words (n) in the new EDP from Pro-Explora is constrained to a minimum
400 of 6 and a maximum of 12 ($6 \leq n \leq 12$). This is based on findings from studies and the
401 result of the EDP word-count analysis on the extracted corpus as shown in Fig. 3.



402

403 **Fig. 3 Word count of titles describing EDP**

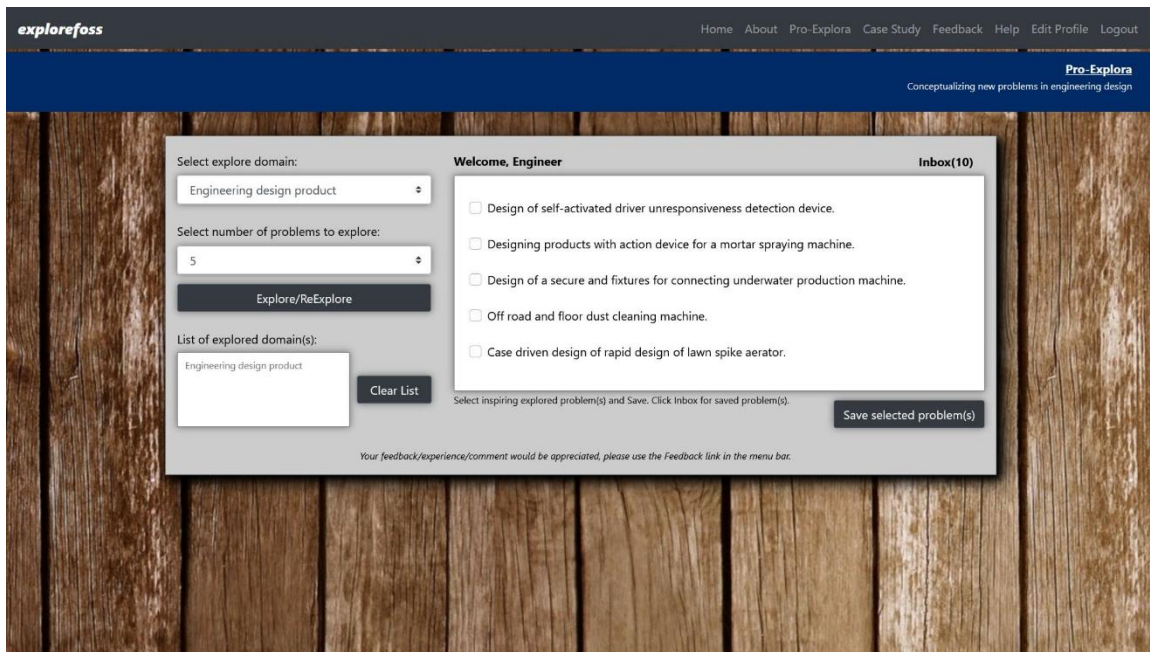
404

405 Shown in Fig. 3 are the word counts mostly used in titles that describe an EDP
406 (Section 3.1.1). It could be seen that the most used word counts range between 8 and 12.
407 This range is significant in this paper and correlates with the scholarly suggestions that a
408 maximum of 12 words should be used to describe an EDP to inspire thoughts and attract

409 attention [80-83]. Also, as discussed in Section 2.3.1, the limit of words the brain can
410 process at once is between 5 and 9. Hence, the word count limit of 6 - 12 for the Pro-
411 Explora output is considered appropriate. Using python code, it is checked that the output
412 from *System B* Fig. 1 ending with a period satisfies the word count limit.

413 Pro-Explora GUI

414 Backend python codes enable the functionality of Pro-Explora. However, it is accessed
415 through a simple web Graphical User Interface (GUI) shown in Fig 4 which requires a Login
416 to access at <https://www.explorefoss.com/>. A logged-in user “Engineer” is shown in the
417 Pro-Explora GUI in Fig. 4. The GUI has two settings that contain some options that could
418 be selected based on preference before EDPE as explained next.



419

420

Fig. 4 Pro-Explora GUI with some framed design problems

421

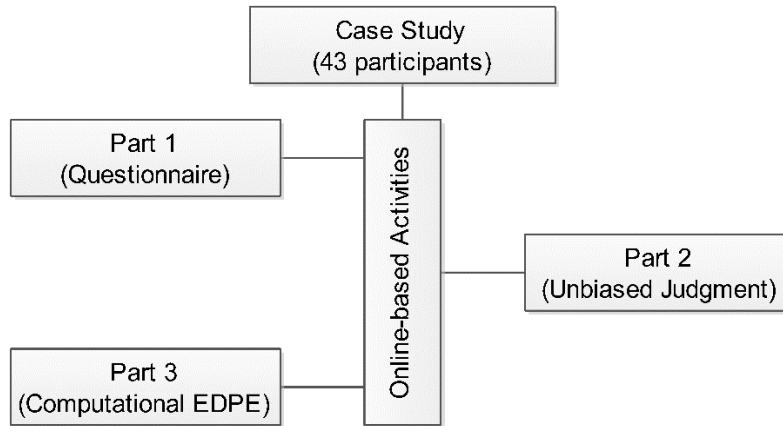
422 The Pro-Explora GUI, as shown in Fig.4, has the “Select explore domain” and “Select
423 number of problems to explore” settings. The “Select explore domain” setting has four
424 domain options – “Engineering design product”, “Engineering design research”,
425 “Engineering design machine intelligence”, and “Engineering design cross-domain”. The
426 “Engineering design cross-domain” option has the largest database, a combination of the
427 other three domains. The “Select number of problems to explore” setting has 1 and 5 as
428 the minimum and maximum numbers to select with 1 as the default. The history of all
429 explored domains is displayed in the “List of explored domain(s)” section of the GUI. With
430 the preferential options set, clicking on the Explore/ReExplore button activates a new
431 EDPE process. The GUI in Fig. 4 displays five framed EDPs and any or all could be selected
432 and saved for later review. Pro-Explora can frame over 100 unique EDPs per minute. This
433 would be impossible with the natural EDPE process (Section 2.1), considering the
434 associated challenges (Section 2.2). The framework and tool for computational support in
435 EDPE are presented in this section. Presented in the next section are the methods of data
436 collection and analyses to answer the *RQs* in this paper.

437

438

439 **4 DATA COLLECTION METHOD AND ANALYSES**

440 **4.1 Case study details**



441

442 **Fig. 5 Overview of data collection approach**

443 As shown in Fig. 5, the data collection method is through a case study comprising
444 three parts. Adverts for the case study are sent through channels targeting professionals
445 and students (Year 2 upwards) within engineering design or design-related academic
446 programmes. Participants for the case study include students from UK and Singapore
447 Universities and professionals from countries including Canada, France, India, Russia,
448 Singapore, the United Kingdom, and the United States. The case study advert specifies
449 eligibility for participation - healthiness, a basic understanding of creativity, and an
450 engineering design-related background. All participants provide qualifications, years of
451 experience, and other details (Table 2). Only one participant who showed interest is
452 disqualified based on not meeting the criteria of educational background. The creativity
453 “understanding” in the eligibility is important because creativity as a phenomenon for
454 coming up with something new and useful is strongly correlated with EDPE [84-88].

455

456

Table 2 Case study participants' detail

Participants	Gender			Academic Qualification (Obtained/in view)			Experience (Years)	
	Male	Female	Total	B	MS	PhD	≤ 3	> 3
Novice	12	7	19	5	11	3	19	0
Experienced	23	1	24	10	9	5	0	24
Total	35	8	43	15	20	8	19	24

457

B – Bachelors, MS – Masters

458

459 As shown in Table 2, the participants are categorized based on their years of
 460 experience as either “Novice” (Mean (M) = 1.5 years, Standard Deviation (SD) = 0.7 years)
 461 or “Experienced” (M = 8 years, SD = 3.8 years, Range (R) = 17 years). An inclusion criterion
 462 of 0 – 40 years of experience is applied to ensure that participants find it easy to perform
 463 computer-based activities. The three parts of the case study shown in Fig. 5 lasted about
 464 30 minutes in total and are presented next.

465

466 **Case Study: Part 1 - Questionnaire responses**

467 This part of the case study addresses *RQ1*. The aim is to subtly test the consciousness of
 468 EDPE practice within the engineering design community. As previously mentioned,
 469 creativity correlates strongly with EDPE and EDPS. The participants are given the
 470 questions in Table 3 as an online questionnaire to respond to. In Table 3, Questions ‘a’,
 471 ‘c’, and ‘d’ are subtly designed to test if participants have a conscious understanding
 472 (personally or taught) of creativity relative to EDPE or identifying a new EDP. An effort is
 473 made to “avoid leading questions” by not mentioning EDPS or EDPE [89].

474

475

Table 3 Questionnaire for Part 1 of the case study

Questionnaire themes
(a) What does it mean to be creative?
(b) What are the major roles of creativity in engineering design?
(c) Do You Consider Yourself Creative?
(d) Why are you creative or uncreative?
(e) Were you taught creativity at University or at work?

476

477

Case Study: Part 2 - Unbiased judgment of natural and computational EDP

478

In this part of the case study, the participants are given the following instruction:

479

“Below, you are presented with 20 unique engineering design-related problems.

480

Professionals within the engineering design discipline conceptualized and produced some

481

of these problems while some are computationally generated with a computational tool.

482

You are required to go through each of the 20 problems and choose one option under

483

each problem based on whether you think the problem is produced by a person or

484

computationally generated.” Two options are provided for response.

485

Different sets of 20 EDPs similar to that in Table 4 are presented to each

486

participant. However, each set contains a randomly arranged 5 EDPs framed by a design

487

engineer (naturally framed) and 15 EDPs framed using Pro-Explora (computationally

488

framed). The ratio (1:3) of the EDPs is intentionally not disclosed to the participants. Since

489

the participants are unaware of this ratio, it helps to eliminate bias in their judgements.

490

To make the reader guess, the categories of the EDPs – “naturally” or “computationally”

491

framed, are not indicated here but in Section 5.4.

492

493

494 **Table 4 A sample set of 20 EDPs for participants**

Naturally and computationally framed EDPs	
495	<ol style="list-style-type: none"> 1. Design of a mechanical intrusive force detection device. 2. To design a portable water distillation device. 3. A sustainable packaging design for wine. 4. Designing an interactive interface for collaborative engineering design. 5. A design of an automatic bottle opener. 6. Towards intelligent emotion detection system for video traffic surveillance. 7. Ai-based learning models for video traffic surveillance. 8. Design and material properties to minimize biofilm deposits. 9. Design of human-powered hybrid electric-power shovel for the physically challenged. 10. Design of self-reconfigurable production equipment during operation. 11. Anti riot drone without traffic lights. 12. Investigation of anomaly detection in a critical materials. 13. Design of a self-timing solar seawater desalination machine. 14. Staging co-design for reverse modeling of product development. 15. Detecting aggressive driving behavior using scilab. 16. Design of remote intelligent home finance software. 17. Designing products by artificial intelligence design approach. 18. A computationally efficient real-time vehicle and speed detection using federated learning. 19. Automatic mechanical footstep power tiller machine. 20. Design of production information retrieval system.

495

496 **Case Study: Part 3 - Evaluating the value of a computational EDPE support tool**

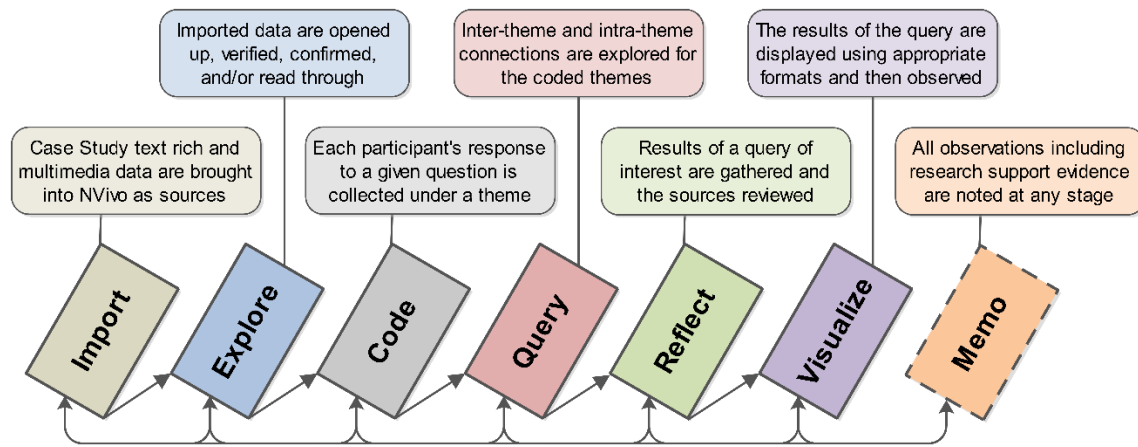
497 This is the last part of the case study and contributes to answering *RQ2*. It is about
 498 evaluating a computational EDPE support tool – Pro-Explora, presented in Section 3.3.
 499 Participants use the tool to come up with at least 5 EDPs in about 10 minutes. On a Likert
 500 scale of 1 – 10, the participants rated the usefulness of the EDP framed by the tool. They
 501 also provide additional information on 1) the reason for the usefulness rating they
 502 provide, and 2) whether the EDP inspired or prompted them to think of a different EDP
 503 related or unrelated to the originally framed EDP.

504

505 4.2 Data analysis

506 Data from the case study is qualitatively and quantitatively analyzed. The
507 qualitative analysis is performed with NVivo 12 - powerful software for qualitative data
508 analysis, following the workflow in Fig. 6.

509



510

511

Fig. 6 Case study qualitative analysis workflow

512 In Fig. 6, the 7 stages involved in qualitative analysis with NVivo are shown. The
513 various data collected are imported into NVivo and arranged. The data is coded - the
514 process of gathering materials (participants' responses) by topics or themes. This is
515 followed by querying the data for patterns and connections. The query results are
516 reflected upon and visualized. Although the workflow in Fig. 6 is iterative, the first 6 stages
517 should be sequentially completed before any iterative update can be made to any of
518 them. However, the last stage (Memo) can be referenced from any stage at any time.

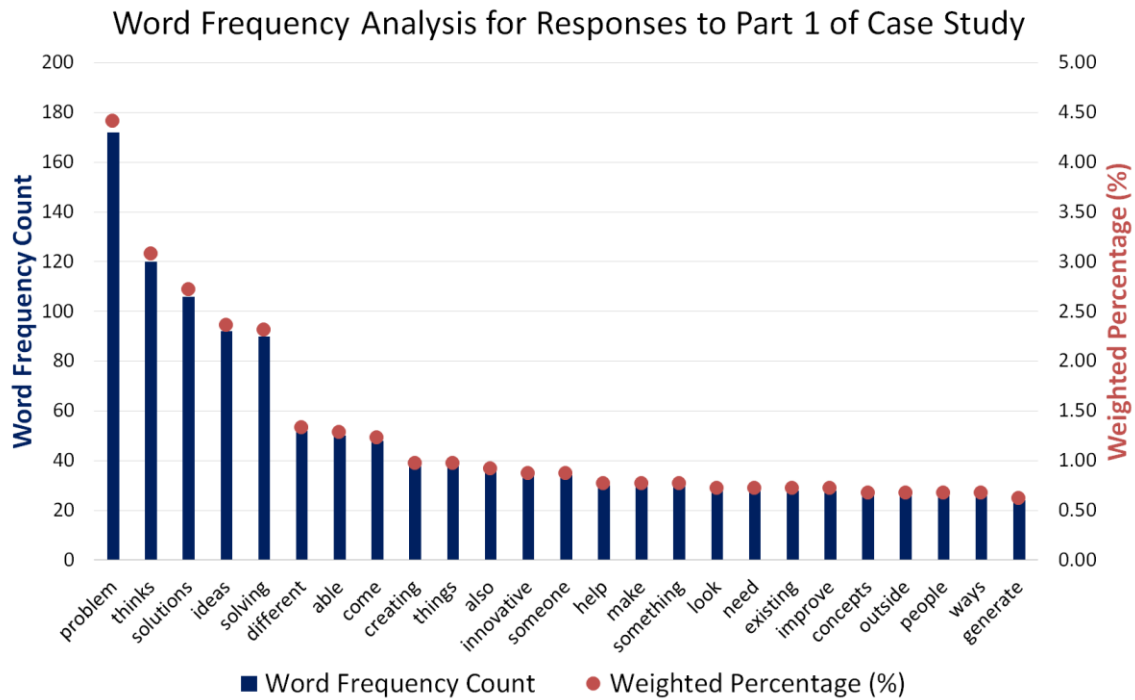
519

520

521 **5 RESULTS**

522 **5.2 General understanding of creativity**

523 In Part 1 of the case study (Section 4.1), Questions ‘a’, ‘b’, and ‘d’ in Table 3 is to
 524 address RQ1. They are designed to subtly reveal the participants’ general understanding
 525 of creativity relative to EDPE, which strongly correlates with creativity as mentioned in
 526 Section 4.1. Responses to the questions are qualitatively analyzed using the WFA in Fig.
 527 7, Cluster Analysis (CA) in Fig. 8, and text query search (Fig. 9.)



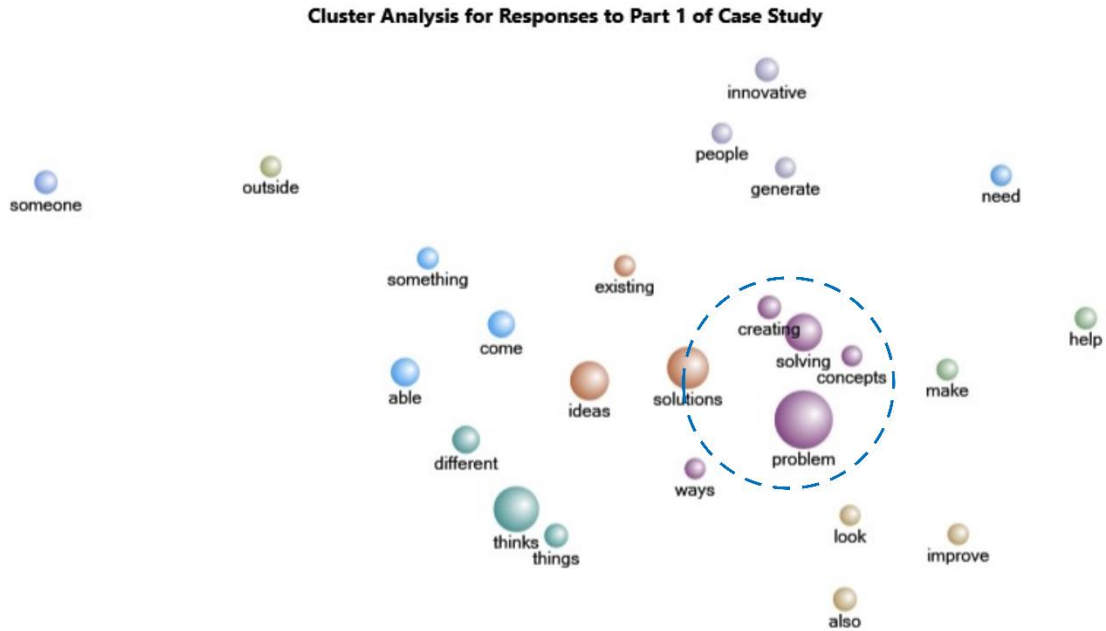
528

529 **Fig. 7 Most frequent words in explaining creativity in engineering design**

530

531 Presented in Fig. 7 are the 25 most frequent words used by the participants to
 532 explain what creativity means to them. It could be seen that the participants mainly
 associate the word ‘problem’ with creativity in engineering design. A cluster analysis (CA)

533 is performed to observe the relationship between the 25 words in Fig. 7. The CA result is
534 shown in Fig. 8.



535

536 **Fig. 8 Word cluster showing a relationship in words explaining creativity**

537

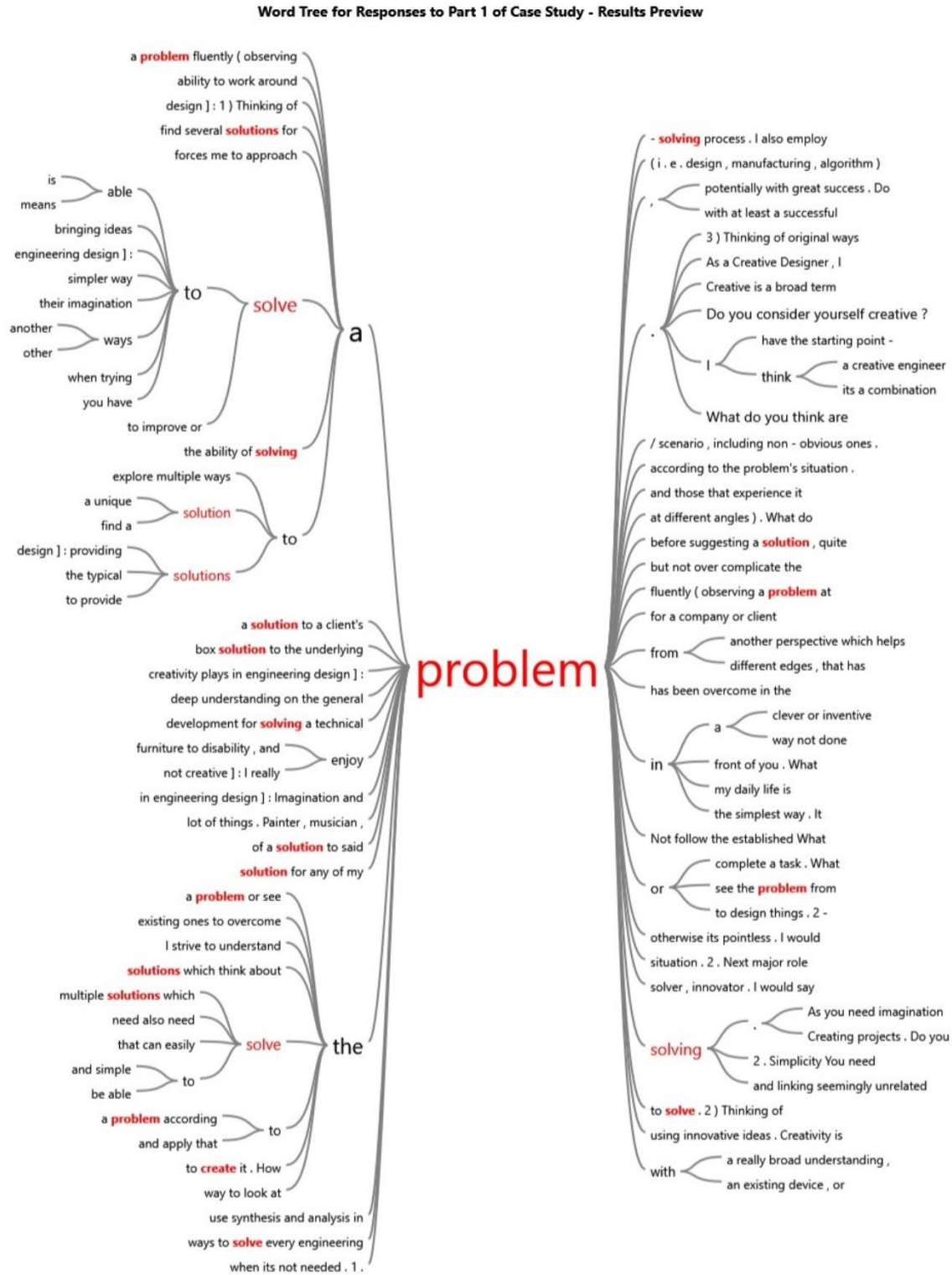
538 Shown in Fig. 8, are clusters suggesting the relationships between the 25 most
539 frequent words in Fig. 7. The highest single cluster of 5 words – ‘concepts’, ‘creating’,
540 ‘solutions’, ‘solving’, and ‘problem’ could be seen encircled in Fig. 8. To understand the
541 context of the encircled cluster, a text query search is run with the 5 cluster words, and
542 the result is presented in Fig. 9.

543

544

545

546



547

548

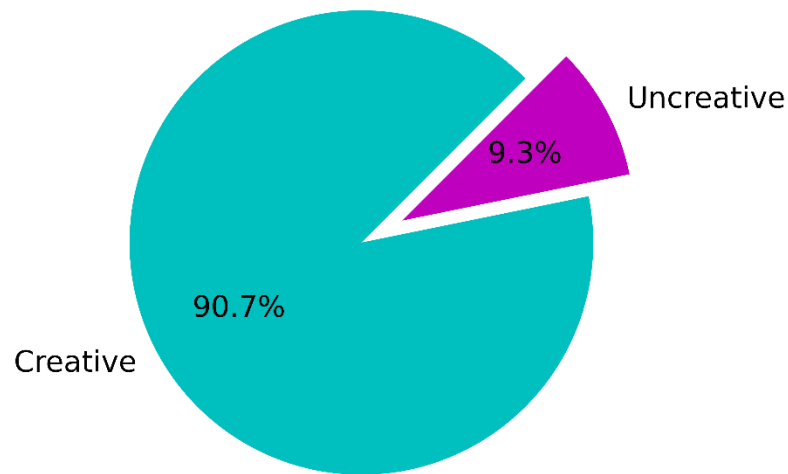
Fig. 9 Word Tree result for a text query search

549

550 The Word Tree (WT) in Fig. 9 shows the root term as ‘problem’ which is the most
551 frequent word in the WFA in Fig. 7. For a clearer context of the relationships in the cluster
552 words in Fig. 8, five words are allowed on either side of the root term in Fig. 9. As the WT
553 shows, creativity is generally understood to be an EDPS phenomenon. This is likely to be
554 the participants' understanding of creativity from academia and literature. There is no
555 explicit association of creativity to EDPE by the participants. The next result presented is
556 on the teaching of creativity in engineering design in academia.

557 **5.3 Teaching creativity in engineering design**

Participants Who Consider Themselves To Be Creative or Uncreative



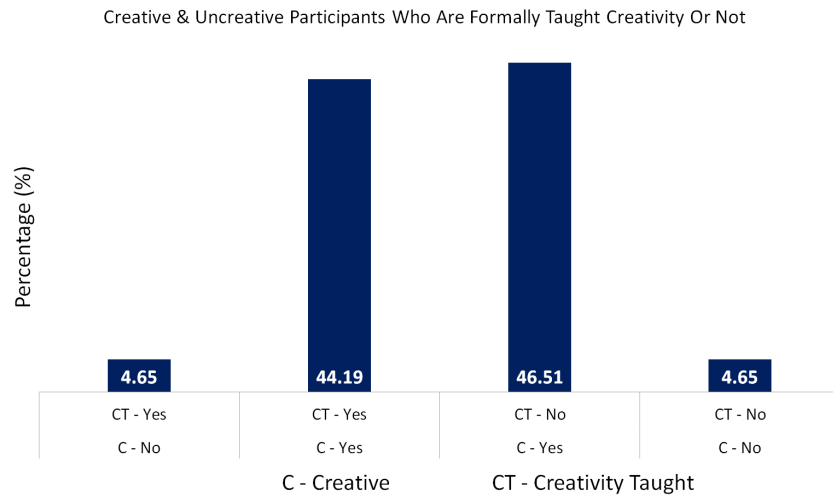
558

559 **Fig. 10 Participants' responses to their creativity ability**

560

561 The participants' responses to Questions 'c' and 'e' of Part 1 of the case study
562 (Table 3) are presented in Fig. 10 and Fig. 11. Questions 'c' and 'e' are designed to reveal

563 the adequacy and focus on creativity teaching in academia. As shown in Fig. 10, most of
 564 the participants consider themselves creative which shows that creativity is a popular skill
 565 in engineering design.



566

Fig. 11 An insight into creativity teaching in academia

567

568 Presented in Fig. 11, are the percentages of participants who are either taught
 569 creativity in academia or industry and those who are not. It could be seen that not all the
 570 participants who consider themselves creative (in Fig. 10) are formally taught creativity.
 571 Some design engineers could be naturally creative without being formally taught as
 572 shown in Fig. 11. However, this should not deter effort in teaching creativity techniques
 573 and skills formally in academia and industry. Every good natural ability needs formal
 574 support. For example, some people are naturally good at playing football but football
 575 academies exist. There is the possibility that those who are creative (C – Yes) but not
 576 taught creativity (CT – No) could have been more creative if formally taught creativity in
 577 academia. Also, the possibility exists that those who are not creative (C – No) and not
 578 taught creativity (CT – No) could have been creative if formally taught. The result

579 presented in Fig. 11 suggests that creativity teaching in academia may be below average
 580 as over 50% of the participants are not formally taught creativity. For the lesser
 581 percentage that is formally taught creativity, the focus is on EDPS while EDPE is ignored
 582 as shown in Fig. 9. Following the completion of Part 1 of the case study, Part 2 is
 583 commenced and the results are presented next.

584

585 **5.4 Differentiating a computationally and naturally framed EDP**

586 In Part 2 of the case study, as part of answering *RQ2*, the intent is to test if the participants
 587 could differentiate between a computationally and naturally framed EDPs. In the set of
 588 20 EDPs presented in Table 4, EDPs 1 – 5 are framed by a design engineer while EDPs 6 –
 589 20 are framed by Pro-Explora. Participants are required to distinguish both categories of
 590 EDPs, as explained in Section 4.1. The result of this activity for all the participants (Novice
 591 and Experienced) is presented in Fig. 12. The “Novice” and “Experienced” participants are
 592 code-named and presented in Table 5 for confidentiality.

593

594

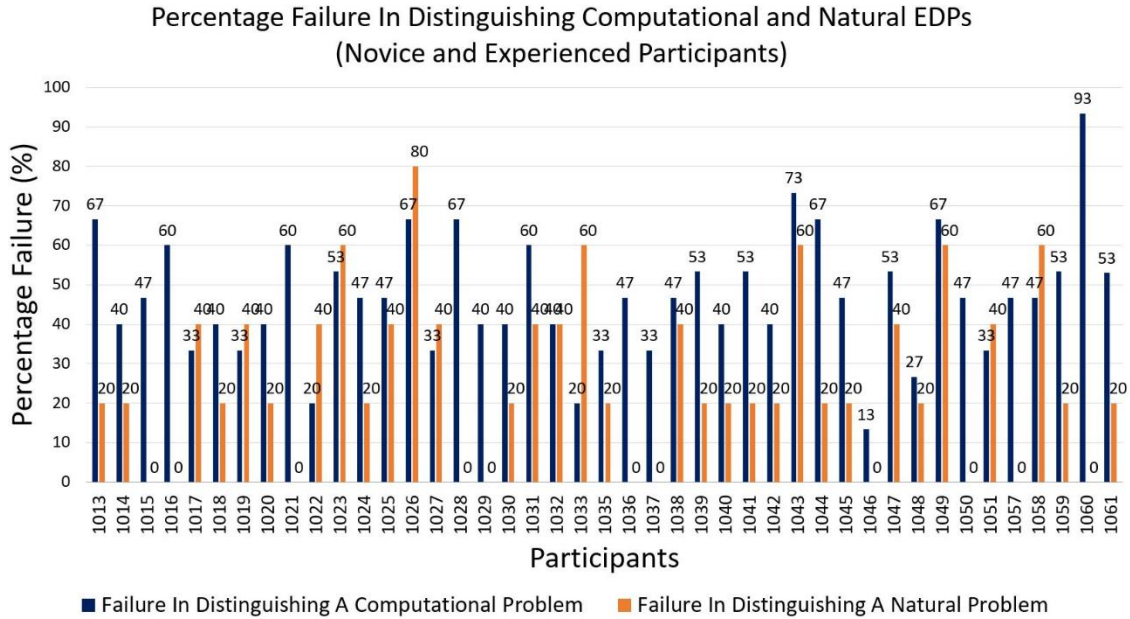
Table 5 Novice and Experienced participants

Novice				Experienced			
1013	1019	1026	1058	1022	1035	1041	1048
1014	1020	1027		1028	1036	1042	1049
1015	1021	1029		1030	1037	1044	1057
1016	1023	1043		1031	1038	1045	1059
1017	1024	1050		1032	1039	1046	1060
1018	1025	1051		1033	1040	1047	1061

595

596

597



598

599

Fig. 12 Result of Part 2 (Unbiased Judgement) of the case study

600

601

602

603

604

605

606

607

608

609

610

611

As shown in Fig. 12, some of the participants have zero failures in distinguishing a naturally framed EDPs. These participants are “1015”, “1016”, “1021”, “1028”, “1029”, “1036”, “1037”, “1046”, “1050”, “1057”, and “1060”. It could be seen in Table 5 that some of these participants are “Novice” while some are “Experienced”. These zero failures suggest that the “computationally framed EDP” judged by the respective participants as a “naturally framed EDP” appears natural, useful, and meaningful. The correlation between the participants’ years of experience and failures in distinguishing a computationally and naturally framed EDP is tested for statistical significance. The results are presented in Table 6. Note that the participants are already categorized based on years of experience in Table 2. Hence, the failures in Table 6 are relative to the participants’ (Novice and Experienced) years of experience.

612 **Table 6 Correlation between experience and distinguishing a computational EDP**
 613 (Null hypothesis (H_0): There is a significant relationship between the participants' experience and
 614 their failures in distinguishing a computationally and naturally framed EDP)

	P-value	Pearson's r
Failures in distinguishing a computationally framed EDP	0.78	0.04
Failures in distinguishing a naturally framed EDP	0.72	-0.06

615 P-value <0.05 defines statistical significance

616

617 **Cosine similarity assessment**

618 The result presented in Fig. 12 indicates a misjudgment of at least one naturally framed
 619 EDP or computationally framed EDP by all the participants. This suggests a similarity
 620 between the two categories of EDPs. As a further analysis, the naturally framed (EDP1 –
 621 EDP5) and computationally framed (EDP6 – EDP10) EDPs in Table 4 are assessed for
 622 differences or similarities. The first EDP in Table 4 is named correspondingly as EDP1, the
 623 second EDP2, and the tenth EDP10. A random quote (Q^*) is added in Table 7 to see how
 624 its similarity compares with the EDPs. The result of the assessment is presented in Table
 625 7. The assessment is performed using cosine similarity, which measures similarity
 626 between texts by “calculating the cosine of the angle between the two vectors” [90]. A
 627 web text trained Spacy pipeline, `en_core_web_lg`, is used to compute the cosine
 628 similarity. Spacy is an open-source python library for NLP. Cosine similarity ranges from 0
 629 – 1 with 1 indicating 100% similarity. It could be seen in Table 7 that most similarities
 630 between the naturally and computationally framed EDPs are above 65%. This justifies the
 631 failures in judgments in Fig. 12. Since Q^* in Table 7 is a quote, its similarity with the EDPs
 632 is the lowest across rows.

633

634

Table 7 Cosine similarity assessment result

	EDP6	EDP7	EDP8	EDP9	EDP10	Q*
EDP1	0.8163	0.7000	0.7461	0.8465	0.8589	0.6163
EDP2	0.7164	0.6240	0.7838	0.6794	0.6686	0.5628
EDP3	0.6700	0.6656	0.6258	0.6776	0.6872	0.5249
EDP4	0.7842	0.7275	0.7191	0.7500	0.7710	0.5371
EDP5	0.6163	0.5767	0.5481	0.7164	0.7306	0.4750

635

Q* - "Anyone who has never made a mistake has never tried anything new." (Albert Einstein)

636

637

5.5 The value of computational support tool in EDPE

638

To address *RQ2*, in Part 3 of the case study, the participants used Pro-Explora as a support

639

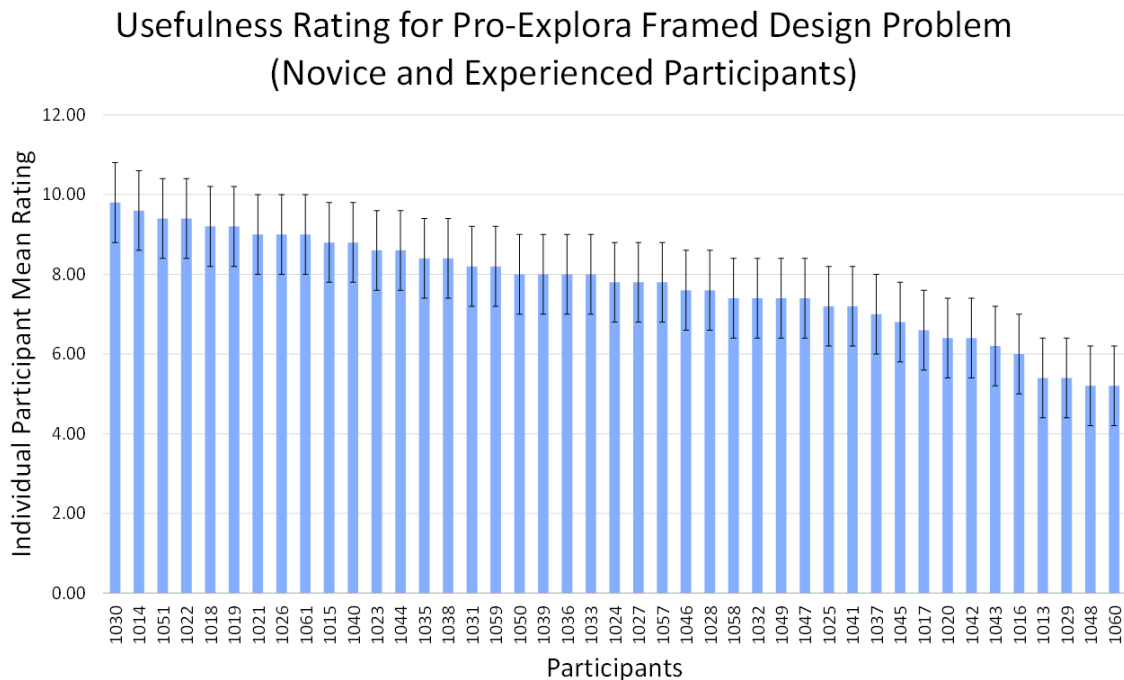
tool to come up with some EDPs. They rated 5 of the EDP on a Likert scale of 1 – 10 (with

640

10 being the highest). In Fig. 13, the mean of the ratings for all participants (Novice and

641

Experienced) is shown with the standard error (SE).



642

643

Fig. 13 Participants' usefulness rating for Pro-Explora framed EDP (with ±1 SE bar)

644 The overall mean usefulness rating of the participants (Novice and Experienced)
645 shown in Fig.13 is 7.74. Coincidentally, the separate mean usefulness ratings of the
646 “Novice” and “Experienced” participants is 7.74 and 7.74, respectively. The mean for a
647 Likert scale of 1 – 10 is 5.5. Hence, the overall usefulness rating of the participants (7.74)
648 for Pro-Explora generated EDPs is above the mean value (5.5) of the Likert scale. This
649 usefulness rating of 7.74 out of a maximum of 10 on the Likert scale could be considered
650 high. In Fig. 13, the error bar overlaps give a visual insight on the variabilities in the
651 individual rating values which aggregate to the means used for the plot. For example, in
652 Fig. 13, the error bar for participant “1030” overlaps with others to the right of participant
653 “1030” up to participant “1033”. This indicates that some of the separate ratings of the
654 participants to the right of “1030” are higher than that of participant “1030”. As seen in
655 Table 5, these participants belong to either the “Novice” or “Experienced” category.

656 Some of the participants mention that they are inspired or prompted to think of a
657 different or related EDP based on the EDP framed by Pro-Explora. The correlation
658 between the participants' experience and their usefulness ratings is statistically analyzed
659 and presented in Table 8. The analyses in Table 8 are relative to the participants' years of
660 experience, as indicated in Table 2. As shown in Table 8, the analysis is performed for the
661 overall participants (Novice and Experienced). As a further confirmation, the analysis is
662 also performed separately for only the “Novice” and only the “Experienced” participants.
663 Coincidentally, as shown in Table 8, the p-value for the overall rating is the same as that
664 of the “Experienced” participants.

665

666 **Table 8 Correlation between experience and rating of a computational EDP**
 667 (Null hypothesis (*H₀*): There is a significant relationship between the participants' level of
 668 experience and their usefulness ratings of Pro-Explora framed EDP)

	P-value	Pearson's r
Usefulness Rating (Overall)	0.69	-0.06
Usefulness Rating (Novice participants only)	0.16	-0.34
Usefulness Rating (Experienced participants only)	0.69	-0.08

669 P-value <0.05 defines statistical significance

670

671 **6 DISCUSSION**

672 **6.1 Academic implications**

673 The findings and results contribute to knowledge by providing empirical evidence
 674 on the 1) lack of focus on EDPE within the engineering design community and 2) value of
 675 computational support in the EDPE process for the first time. The lack of attention on
 676 EDPE contrasts with the standard expectation of design engineers in identifying societal
 677 EDP using their experience, knowledge, and background [4, 10, 13, 91]. The natural EDPE
 678 process investigated in this paper requires creativity. Over 50% of the participants in the
 679 case study indicate that they were not formally being taught creativity. This suggests a
 680 lack of creativity teaching in academia within engineering design disciplines [92]. Also, the
 681 effort in teaching creativity in academia is focused on EDPS while EDPE is ignored. The
 682 case study result indicates that the general understanding of creativity is about EDPS
 683 within the engineering design community. This understanding is likely from the teachings
 684 provided in academia. Hence, effort in teaching creativity in engineering design disciplines
 685 should equally focus on both EDPS and EDPE.

686

687 **6.2 Industry implications**

688 There are scholarly opinions, as mentioned previously, that EDPE is a challenging
689 activity. However, a paper on why EDPE is challenging lacks. This paper highlights the
690 possible determinants of the challenges associated with EDPE. Hence, this makes it
691 possible to extensively investigate some computational technologies that could support
692 the natural EDPE process, while it was previously indicated that computational EDPE
693 would be impossible [93]. This paper would provide opportunities for further research in
694 the area of computational EDPE in engineering design and other fields. For example, it
695 could be applied in the medical field to identify new possibilities. The results presented in
696 this paper show that both novice and experienced design engineers can come up with at
697 least 5 EDPs in about 10 minutes. This feat would be difficult or impossible within natural
698 limits due to cognitive limitations and fatigue (Section 2.2). In the natural EDPE process,
699 it is considered “entirely reasonable to spend several months or longer thinking about
700 potential problems” to solve [94]. The EDP framed by Pro-Explora is given an average
701 “usefulness” rating of 7.74 out of 10 by both novice and experienced design engineers
702 (Section 5.5). This indicates that design engineers could be computationally and
703 intentionally inspired, prompted, or supported in using their knowledge in EDPE. The
704 inspiration occurs when the Pro-Explora framed EDP coincides with the design engineer’s
705 knowledge, experience, and/or background. This is similar to serendipity discovery
706 (Section 2.1), and some participants agree that the EDP framed by Pro-Explora inspired
707 them to think of a different EDP. Knowledge is infinite, and design engineers cannot
708 measure their knowledge or intentionally recall all they know [95, 96]. Hence,

709 computationally prompting the design engineer of an EDP that may be within the domain
710 of their knowledge to solve is advantageous and a rapid way of discovering a new EDP, an
711 invention, or innovation.

712

713 **6.3 Limitations and opportunities**

714 The results are based on the direct responses provided by the participants. No further
715 verification of the information is carried out. For example, the Universities attended by
716 the participants who reported that they are not taught creativity are not contacted for
717 verification. Also, being an online activity, it is not certain whether the participants spent
718 longer or lesser than 10 minutes during EDPE with Pro-Explora. However, an instruction
719 to spend 10 minutes on the task is provided. During the case study, the participants used
720 Pro-Explora once for EDPE, rated its outputs above average, and requested access for
721 continued use which is granted. Further trials would be necessary to monitor the
722 subsequent rating for Pro-Explora and ensure an increased rating. The uniqueness of Pro-
723 Explora framed EDP is based on a duplication recognition search in the original corpus
724 used in generating the EDP. This search is not extended to the google and patent
725 databases which are popular for verifying uniqueness. However, during a pilot test, a
726 manual search on google returned no duplicate for any Pro-Explora framed EDP.

727 Although BERT and LSTM technologies are potential computational technologies
728 for EDPE (Section 2.3), they have not been used to compare with the MM used in this
729 paper. Being in its infancy (Version 1), Pro-Explora will be improved further based on the
730 feedback received from the participants. This will include optimizing its outputs and

731 exploring other related NLP technologies including BERT and LSTM. Data collection for
732 Pro-Explora database will continue, and its model will be updated continuously.

733

734 **7 CONCLUSIONS**

735 In this paper, case study-based evidence is provided to highlight the lack of
736 attention on EDPE - an important aspect of engineering design at early-stage product
737 design and development. Albeit there are few studies on the lack of attention on EDPE, a
738 study providing empirical evidence and determinants for it lacks. The natural approaches
739 related to EDPE are investigated including the “garbage can” model and serendipity
740 phenomenon. Some challenges and natural limitations associated with the natural EDPE
741 approach are identified including cognitive fatigue. This suggests that computational
742 support could be advantageous in the process. In response, a data-driven computational
743 EDPE framework and support tool – Pro-Explora are presented. The tool is the first-of-its-
744 kind computational technology that mimics the natural EDPE process. It is based on a
745 synergy of the MM and some big data technologies including ML and NLP.

746 A case study is conducted with 43 participants including novice and experienced
747 design engineers. During the case study, the participants could not distinguish EDP
748 framed by Pro-Explora when presented alongside naturally framed ones. Using Pro-
749 Explora as support, novice and experienced participants come up with at least 5 new EDPs
750 in about 10 minutes. This would be difficult or impossible with the natural EDPE approach.
751 The overall average rating provided by the participants on the usefulness of Pro-Explora
752 framed EDP is 7.74 out of 10. This is promising for accelerated innovations and inventions

753 in the industry. Further, the result shows that over 50% of the participants in the case
754 study did not receive any formal teaching on creativity in academia. This highlights the
755 importance of focusing on teaching creativity in engineering design-related disciplines
756 which is fundamental in EDPE.

757

758 **ACKNOWLEDGMENT**

759 All the students and professionals including academics and inventors who contributed to
760 this study by participating in the case study are hereby acknowledged.

761

762 **REFERENCES**

763

- 764 [1] Olewnik, A., Yerrick, R., Simmons, A., Lee, Y., and Stuhlmiller, B., 2020, "Defining
765 open-ended problem solving through problem typology framework," *International*
766 *Journal of Engineering Pedagogy (IJEP)*, 10(1).
- 767 [2] Stevens, R., Johri, A., and O'connor, K., 2014, "Professional engineering work,"
768 *Cambridge handbook of engineering education research*, pp. 119-137.
- 769 [3] Vale, J., Gordon, K., Kirkscey, R., and Hill, J., 2020, "Student and Faculty
770 Perceptions of Capstone Purposes: What can Engineering Learn from Other
771 Disciplines?," *Proceedings of the Canadian Engineering Education Association (CEEA)*.
- 772 [4] Taura, T., 2015, *Principia Designae-Pre-Design, Design, and Post-Design: Social*
773 *Motive for the Highly Advanced Technological Society*, Springer.
- 774 [5] Walter, C., and Richards, E., 1996, "Engineering and the law [custodians of
775 information]," *IEEE Engineering in Medicine and Biology Magazine*, 15(5), pp. 138-139.
- 776 [6] Jaussi, K. S., and Topaloglu, E., 2020, "Intentionality," *Encyclopedia of Creativity*
777 *(Third Edition)*, S. Pritzker, and M. Runco, eds., Academic Press, Oxford, pp. 672-677.
- 778 [7] Getzels, J. W., 1979, "Problem finding: A theoretical note," *Cognitive Science - A*
779 *Multidisciplinary Journal*, 3(2), pp. 167-172.
- 780 [8] Obieke, C. C., Milisavljevic-Syed, J., and Han, J., 2021, "Data-driven creativity:
781 computational problem-exploring in engineering design," *Proceedings of the Design*
782 *Society*, 1, pp. 831-840.
- 783 [9] NRC - National Research Council, 2012, *A framework for K-12 science education:*
784 *Practices, crosscutting concepts, and core ideas*, National Academies Press.
- 785 [10] NAE - National Academy of Engineering, U. S., 2004, *The engineer of 2020:*
786 *Visions of engineering in the new century*, National Academies Press Washington, DC.

- 787 [11] Obieike, C., Milisavljevic-Syed, J., and Han, J., 2020, "Supporting Design Problem-
788 exploring with Emergent Technologies," *Procedia CIRP*, 91, pp. 373-381.
- 789 [12] Abdulla, A. M., Paek, S. H., Cramond, B., and Runco, M. A., 2018, "Problem
790 Finding and Creativity: A Meta-Analytic Review," *Psychology of Aesthetics, Creativity,*
791 *and the Arts*, 14(1), pp. 3-14.
- 792 [13] Heimlich, J., Wasserman, D., and Hayde, D., 2014, "Human Plus: Real Lives+ Real
793 Engineering," Center for Research and Evaluation.
- 794 [14] Karwowski, M., Jankowska, D. M., Brzeski, A., Czerwonka, M., Gajda, A., Lebuda,
795 I., and Beghetto, R. A., 2020, "Delving into creativity and learning," *Creativity Research*
796 *Journal*, 32(1), pp. 4-16.
- 797 [15] Design Council, 2018, "The Design Economy 2018: The state of design in the UK,"
798 Design Council.
- 799 [16] Mullen, C. A., 2018, *Creativity Under Duress in Education?: Resistive Theories,*
800 *Practices, and Actions*, Springer.
- 801 [17] AlMaian, R. Y., "Analysis of the stakeholders of engineering education system to
802 improve the creativity of engineering education," *Proc. 2017 IEEE International*
803 *Conference on Industrial Engineering and Engineering Management (IEEM)*, IEEE, pp.
804 110-114.
- 805 [18] Chulvi, V., Sonseca, Á., Mulet, E., and Chakrabarti, A., 2012, "Assessment of the
806 relationships among design methods, design activities, and creativity," *Journal of*
807 *Mechanical Design*, 134(11), p. 11.
- 808 [19] Miller, S. R., Hunter, S. T., Starkey, E., Ramachandran, S., Ahmed, F., and Fuge,
809 M., 2021, "How should we measure creativity in engineering design? a comparison
810 between social science and engineering approaches," *Journal of Mechanical Design*,
811 143(3).
- 812 [20] Charyton, C., 2015, *Creativity and innovation among science and art: A discussion*
813 *of the two cultures*, Springer.
- 814 [21] Gross, D. P., "Hiding in Plain Sight: The Identification and (Humble) Origins of
815 General Purpose Technologies.," "Proc. One Hundred Flows" Conference. All-UC Group
816 in Economic History. Cosponsored by the Berkeley Economic History Lab (BEHL) and
817 the Dept. of Economics, UC Berkeley, Berkeley, CA, USA.
- 818 [22] Shneiderman, B., 2007, "Creativity support tools: Accelerating discovery and
819 innovation," *Communications of the ACM*, 50(12), pp. 20-32.
- 820 [23] Nicholl, B., and McLellan, R., "The Contribution of Product Analysis to Fixation in
821 Students' Design and Technology Work," *Proc. The Design and Technology Association*
822 *International Research Conference 2007*, p. 71.
- 823 [24] Harris, S. D., and Zeisler, S., 2002, "Weak signals: Detecting the next big thing,"
824 *The Futurist*, 36(6), p. 21.
- 825 [25] Fischer, G., 1994, "Turning breakdowns into opportunities for creativity,"
826 *Knowledge-Based Systems*, 7(4), pp. 221-232.
- 827 [26] Einstein, A., and Infeld, L., 1938, *The evolution of physics*, Simon & Schuster, New
828 York.
- 829 [27] Lee, J. W., Daly, S. R., Huang-Saad, A., Rodriguez, G., and Seifert, C. M., 2020,
830 "Cognitive strategies in solution mapping: How engineering designers identify problems
831 for technological solutions," *Design studies*, 71, p. 100967.

- 832 [28] Alter, S., and Dennis, A. R., 2002, "Selecting research topics: Personal experiences
833 and speculations for the future," *Communications of the Association for Information*
834 *Systems*, 8(1), p. 21.
- 835 [29] Dennis, A. R., and Valacich, J. S., 2001, "Conducting experimental research in
836 information systems," *Communications of the association for information systems*, 7(1),
837 p. 5.
- 838 [30] Martin, J., 1982, "A garbage can model of the research process," *Judgment calls in*
839 *research*, pp. 17-40.
- 840 [31] Russ, S. W., and Hoffmann, J. D., 2020, "Associative Theory ☆," *Encyclopedia of*
841 *Creativity (Third Edition)*, S. Pritzker, and M. Runco, eds., Academic Press, Oxford, pp.
842 76-82.
- 843 [32] Banan, S., Ridwan, M., and Adisaputera, A., 2020, "A Study of Connectionism
844 Theory," *Budapest International Research and Critics Institute (BIRCI-Journal):*
845 *Humanities and Social Sciences*, 3(3), pp. 2335-2342.
- 846 [33] Gall, M., Gall, J., and Borg, W., 2007, "edition 8," *Educational research: An*
847 *introduction*. Boston. Pearson/Allyn & Bacon.
- 848 [34] Darbellay, F., 2020, "Serendipity," *Encyclopedia of Creativity (Third Edition)*, S.
849 Pritzker, and M. Runco, eds., Academic Press, Oxford, pp. 470-474.
- 850 [35] Mednick, S., 1962, "The associative basis of the creative process," *Psychological*
851 *review*, 69(3), p. 220.
- 852 [36] Dawes, H., Dawes, C., Martin, G., and Macfarlane, A., 2006, "Making Things from
853 New Ideas," *History of Technology Volume 26, 2005: Including Special Issue:*
854 *Engineering Disasters*, 26, p. 1.
- 855 [37] Ishikawa, A., 2010, "Discovery, invention and serendipity," *Chinese Business*
856 *Review*, 9(11), p. 61.
- 857 [38] Hubbell, M., Hard, S., Boots, M., Clarke, M. A., and Smith, J. E., "Pitch Stability of
858 an Unpowered Ground Effect Vehicle," *Proc. ASME International Mechanical*
859 *Engineering Congress and Exposition*, pp. 191-199.
- 860 [39] Thaler, S., 2020, "The creativity machine paradigm," *Encyclopedia of creativity,*
861 *invention, innovation, and entrepreneurship*, E. G. Carayannis, ed., Springer Nature
862 Switzerland AG, pp. 650 - 658.
- 863 [40] Revell, T., and Andersen, K., 2021, "THINGS: IMAGINING WITH," *Designing*
864 *Smart Objects in Everyday Life: Intelligences, Agencies, Ecologies*, p. 57.
- 865 [41] Heath, S., and Shine, B., 2021, "Teaching Techniques to Facilitate Time
866 Management in Remote and Online Teaching," *Journal of Teaching and Learning with*
867 *Technology*, 10(1).
- 868 [42] Miller, G. A., 1956, "The magical number seven, plus or minus two: Some limits on
869 our capacity for processing information," *Psychological review*, 63(2), p. 81.
- 870 [43] Jones, D. M., "The 7±2 urban legend," *Proc. MISRA C 2002 conference* www.knosof.co.uk/cbook/misart.pdf.
- 871 [44] Marois, R., and Ivanoff, J., 2005, "Capacity limits of information processing in the
872 brain," *Trends in cognitive sciences*, 9(6), pp. 296-305.
- 873 [45] Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., and Roelands,
874 B., 2017, "The effects of mental fatigue on physical performance: a systematic review,"
875 *Sports medicine*, 47(8), pp. 1569-1588.
- 876

- 877 [46] Smith, M. R., Coutts, A. J., Merlini, M., Deprez, D., Lenoir, M., and Marcora, S. M.,
878 2016, "Mental fatigue impairs soccer-specific physical and technical performance," *Med*
879 *Sci Sports Exerc*, 48(2), pp. 267-276.
- 880 [47] Boden, M. A., 2007, "Creativity in a nutshell," *Think*, 5(15), pp. 83-96.
- 881 [48] Hurley, P. J., 2021, "Reconceptualizing Ego Depletion as Transient Cognitive
882 Fatigue," Available at SSRN 3797263.
- 883 [49] Stanton, P., 2018, *Conscious Creativity: Look, Connect, Create*, Leaping Hare Press.
- 884 [50] Ariza-Montes, A., Arjona-Fuentes, J. M., Han, H., and Law, R., 2017, "Employee
885 responsibility and basic human values in the hospitality sector," *International Journal of*
886 *Hospitality Management*, 62, pp. 78-87.
- 887 [51] Marcora, S. M., Staiano, W., and Manning, V., 2009, "Mental fatigue impairs
888 physical performance in humans," *Journal of applied physiology*, 106(3), pp. 857-864.
- 889 [52] Southern, S., and Domzalski, S., "Developing Intuition: The Key to Creative Futures
890 Research," *Proc. Annual Meeting of the American Educational Research Association*, p.
891 105.
- 892 [53] Chakrabarti, A., "Towards a measure for assessing creative influences of a creativity
893 technique," *Proc. DS 31: Proceedings of ICED 03, the 14th International Conference on*
894 *Engineering Design*, Stockholm.
- 895 [54] Sommerville, I., 2011, "Software engineering 9th Edition," ISBN-10, 137035152, p.
896 18.
- 897 [55] Bridle, J., 2018, *New dark age: Technology and the end of the future*, Verso Books.
- 898 [56] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K., 2018, "Bert: Pre-training of
899 deep bidirectional transformers for language understanding," *arXiv preprint*
900 *arXiv:1810.04805*.
- 901 [57] Arras, L., Arjona-Medina, J., Widrich, M., Montavon, G., Gillhofer, M., Müller, K.-
902 R., Hochreiter, S., and Samek, W., 2019, "Explaining and interpreting LSTMs,"
903 *Explainable ai: Interpreting, explaining and visualizing deep learning*, Springer, pp. 211-
904 238.
- 905 [58] Wang, A., and Cho, K., 2019, "Bert has a mouth, and it must speak: Bert as a
906 markov random field language model," *arXiv preprint arXiv:1902.04094*.
- 907 [59] Banerjee, A., Fuchsbaue, G., Peikert, C., Pietrzak, K., and Stevens, S., "Key-
908 homomorphic constrained pseudorandom functions," *Proc. Theory of Cryptography*
909 *Conference*, Springer, pp. 31-60.
- 910 [60] Meyn, S., and Tweedie, R. L., 2009, *Markov Chains and Stochastic Stability*,
911 Cambridge University Press, Cambridge.
- 912 [61] Schurz, G., 1996, "Kinds of unpredictability in deterministic systems," *Law and*
913 *prediction in the light of chaos research*, Springer, pp. 123-141.
- 914 [62] Krishnakumar, S., Berdanier, C., Lauff, C., McComb, C., and Menold, J., 2022,
915 "The story novice designers tell: How rhetorical structures and prototyping shape
916 communication with external audiences," *Design Studies*, 82, p. 101133.
- 917 [63] Mehta, P., Malviya, M., McComb, C., Manogharan, G., and Berdanier, C. G., 2020,
918 "Mining design heuristics for additive manufacturing via eye-tracking methods and
919 hidden Markov modeling," *Journal of Mechanical Design*, 142(12).
- 920 [64] McComb, C., Cagan, J., and Kotovsky, K., 2017, "Capturing human sequence-
921 learning abilities in configuration design tasks through Markov chains," *Journal of*
922 *Mechanical Design*, 139(9).

- 923 [65] Rahman, M. H., Xie, C., and Sha, Z., 2021, "Predicting sequential design decisions
924 using the function-behavior-structure design process model and recurrent neural
925 networks," *Journal of Mechanical Design*, 143(8).
- 926 [66] Anandakumar, H., Arulmurugan, R., and Onn, C. C., 2019, "Big Data Analytics for
927 Sustainable Computing," *Mobile Networks and Applications*, 24(6), pp. 1751-1754.
- 928 [67] Thanaki, J., 2017, *Python natural language processing*, Packt Publishing Ltd.
- 929 [68] IPO - Intellectual Property Office UK, 2017, *IP Basics*, Crown Copyright, Newport.
- 930 [69] Langford, C. A., and Pearce, P. F., 2019, "Increasing visibility for your work: The
931 importance of a well-written title," *American Association of Nurse Practitioners*, 31(4),
932 pp. 217-218.
- 933 [70] Greenspan, Y. F., 2016, *A guide to teaching elementary science: Ten easy steps*,
934 Brill Sense.
- 935 [71] Sri, M., 2021, "NLP in Customer Service," *Practical Natural Language Processing*
936 *with Python*, Springer, pp. 13-63.
- 937 [72] Hajba, G., 2018, "Website Scraping with Python: Using BeautifulSoup."
- 938 [73] Mikes, A., Edmonds, K., Stone, R. B., and DuPont, B., "Optimizing an algorithm for
939 data mining a design repository to automate functional modeling," *Proc. International*
940 *Design Engineering Technical Conferences and Computers and Information in*
941 *Engineering Conference*, American Society of Mechanical Engineers, p. V11AT11A014.
- 942 [74] Plötz, T., and Fink, G. A., 2011, *Markov Models for Handwriting Recognition*,
943 Springer.
- 944 [75] Yadav, P., Ladha, S., Deshpande, S., and Curry, E., "Computational model for urban
945 growth using socioeconomic latent parameters," *Proc. Joint European Conference on*
946 *Machine Learning and Knowledge Discovery in Databases*, Springer, pp. 65-78.
- 947 [76] Rabiner, L. R., 1989, "A tutorial on hidden Markov models and selected applications
948 in speech recognition," *Proceedings of the IEEE*, 77(2), pp. 257-286.
- 949 [77] Privault, N., 2013, *Understanding Markov Chains*, Springer.
- 950 [78] Eagle, A., 2005, "Randomness is unpredictability," *The British Journal for the*
951 *Philosophy of Science*, 56(4), pp. 749-790.
- 952 [79] Sheskin, T. J., 2011, *Markov chains and decision processes for engineers and*
953 *managers*, CRC Press, New York.
- 954 [80] Koronis, G., Casakin, H., Silva, A., and Kang, J. K. S., 2021, "The influence of
955 design brief information on creative outcomes by novice and advanced students,"
956 *Proceedings of the Design Society*, 1, pp. 3041-3050.
- 957 [81] Bahadoran, Z., Mirmiran, P., Kashfi, K., and Ghasemi, A., 2019, "The principles of
958 biomedical scientific writing," *International Journal of Endocrinology and Metabolism*,
959 17(4).
- 960 [82] Awasthi, R., and Kulkarni, G. T., 2014, "It's All In The Title: Writing a Proper Title
961 to the Paper," *Journal of Chronotherapy and Drug Delivery*, 5(3), pp. S13 - S16.
- 962 [83] Hays, J. C., 2010, "Eight recommendations for writing titles of scientific
963 manuscripts," *Public Health Nursing*, 27(2), pp. 101-103.
- 964 [84] Wohl, H., 2022, "Innovation and creativity in creative industries," *Sociology*
965 *Compass*, 16(2), p. e12956.
- 966 [85] Godart, F., Seong, S., and Phillips, D. J., 2020, "The sociology of creativity:
967 Elements, structures, and audiences," *Annual Review of Sociology*, 46, pp. 489-510.

- 968 [86] Han, J., Forbes, H., and Schaefer, D., "An exploration of the relations between
969 functionality, aesthetics and creativity in design," Proc. Proceedings of the Design
970 Society: International Conference on Engineering Design, Cambridge University Press,
971 pp. 259-268.
- 972 [87] Polanyi, M., 1958, *Personal Knowledge: Towards a post critical philosophy*,
973 Routledge & Kegan Paul, London.
- 974 [88] Jørgensen, U., "Engineering design competences—controversial relations between
975 techno-science discipline and engineering practice domains pointing to new foundations
976 for engineering knowledge," Proc. INES Workshop, Virginia Tech, Citeseer, pp. 1-16.
- 977 [89] Malhotra, N. K., 2006, "Questionnaire design and scale development," *The*
978 *handbook of marketing research: Uses, misuses, and future advances*, pp. 83-94.
- 979 [90] Kulkarni, A., and Shivananda, A., 2021, *Natural language processing recipes :
980 unlocking text data with machine learning and deep learning using Python / Akshay
981 Kulkarni, Adarsha Shivananda*, Apress, Berkeley, CA.
- 982 [91] ABET - Accreditation Board for Engineering and Technology, 2019, "Criteria for
983 Accrediting Engineering Programs 2020-2021," Engineering Accreditation Commission,
984 ABET, Baltimore.
- 985 [92] Kazerounian, K., and Foley, S., 2007, "Barriers to creativity in engineering
986 education: A study of instructors and students perceptions," *Journal of Mechanical
987 Design*, 129(7), pp. 761–768.
- 988 [93] Celik, Y., 2019, "Rules for Doing and Managing Research Project,"
989 <https://dergi.biruni.edu.tr/wp-content/uploads/2019/12/2-Yusuf-Celik.pdf>.
- 990 [94] Gall, M. D., Borg, W. R., and Gall, J. P., 1996, *Educational research: An
991 introduction*, Longman Publishing.
- 992 [95] Wood, D., 2013, "Fichtes Conception of Infinity in the Bestimmung des Menschcn,"
993 State University of New York Press, pp. 155-171.
- 994 [96] Saad, I., and Chakhar, S., "Aggregation procedure based on majority principle for
995 collective identification of firm's crucial knowledge," Proc. World Summit on
996 Knowledge Society, Springer, pp. 346-352.
- 997
998
999
1000
1001
1002
1003
1004
1005
1006
1007

1008
1009

Figure Captions List

- Fig. 1 Big data-driven computational EDPE framework
- Fig. 2 A theoretical model for EDP sequencing
- Fig. 3 Word count of titles describing EDP and Non-EDP
- Fig. 4 Pro-Explora GUI with some framed design problems
- Fig. 5 Overview of data collection approach
- Fig. 6 Case study qualitative analysis workflow
- Fig. 7 Most frequent words in explaining creativity in engineering design
- Fig. 8 Word cluster showing a relationship in words explaining creativity
- Fig. 9 Word Tree result for a text query search
- Fig. 10 Participants' responses to their creativity ability
- Fig. 11 An insight into creativity teaching in academia
- Fig. 12 Result of Part 2 (Unbiased Judgement) of the case study
- Fig. 13 Participants' usefulness rating for Pro-Explora framed EDP (with ± 1 SE bar)

1010
1011

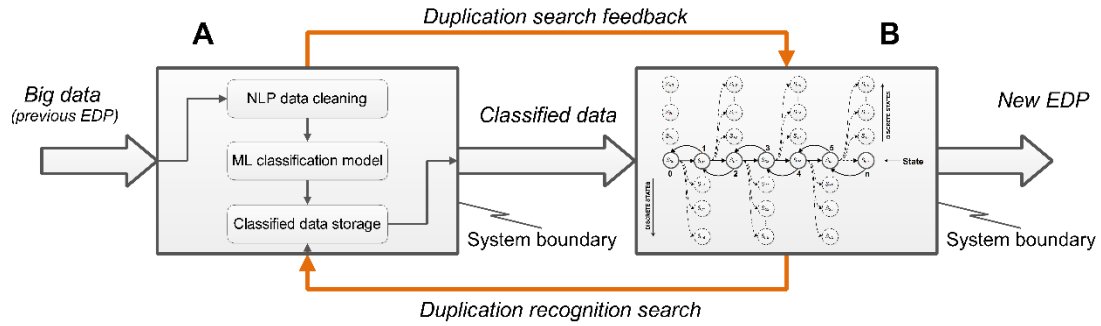
1012
1013

Table Caption List

Table 1	Comparison of algorithms for natural and computational EDPE
Table 2	Case study participants' detail
Table 3	Questionnaire for Part 1 of the case study
Table 4	A sample set of 20 EDP for participants
Table 5	Novice and Experienced participants
Table 6	Relationship between experience and distinguishing a computational EDP
Table 7	Cosine similarity assessment result
Table 8	Relationship between experience and rating of a computational EDP

1014

1015



1016

1017

1018

Fig. 1 Big data-driven computational EDPE framework

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

1042

1043

1044

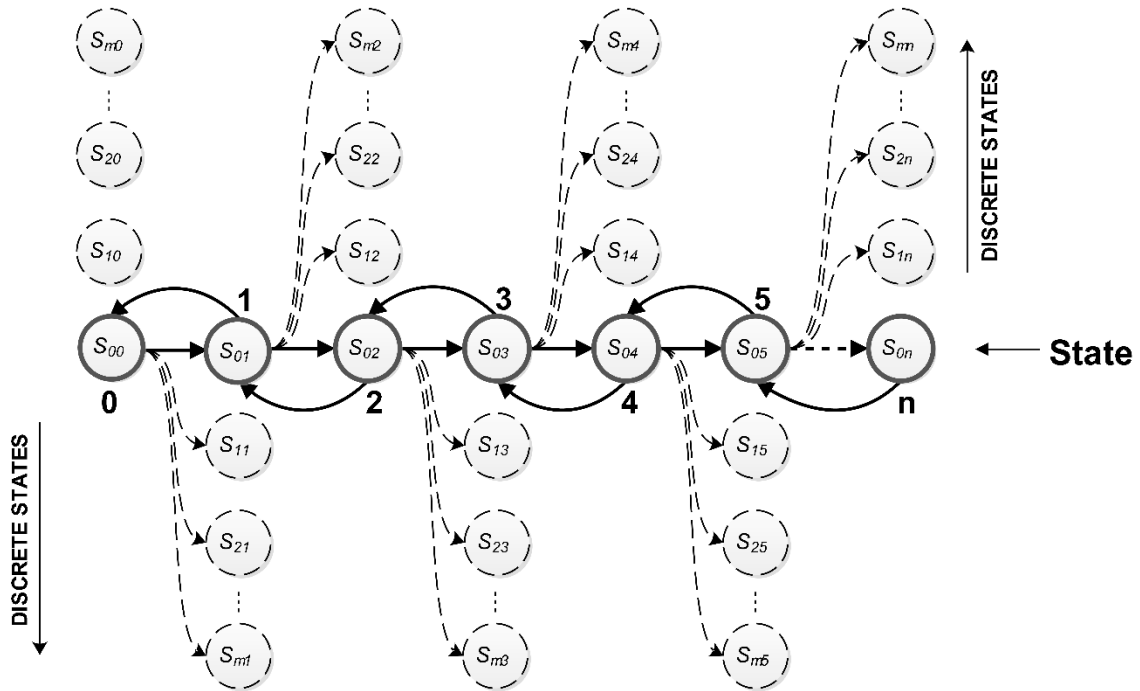
1045

1046

1047

1048

1049



1050
1051

1052

Fig. 2 A theoretical model for EDP sequencing

1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073

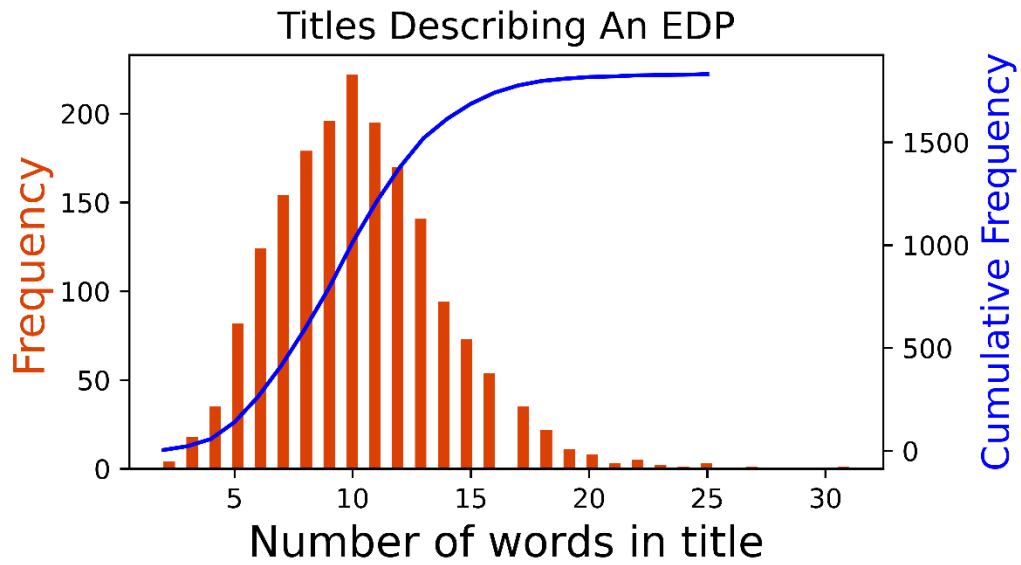
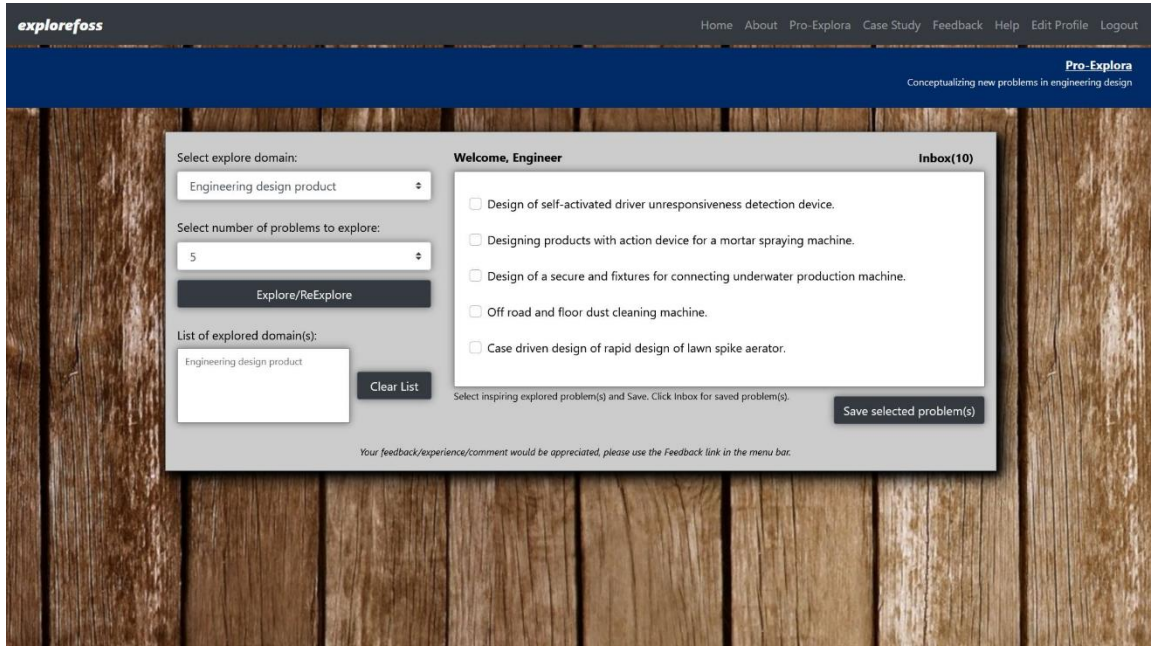


Fig. 3 Word count of titles describing EDP and Non-EDP

1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098

1099
1100
1101
1102
1103



1104
1105
1106
1107

Fig. 4 Pro-Explora GUI with some framed design problems

1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124

1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158

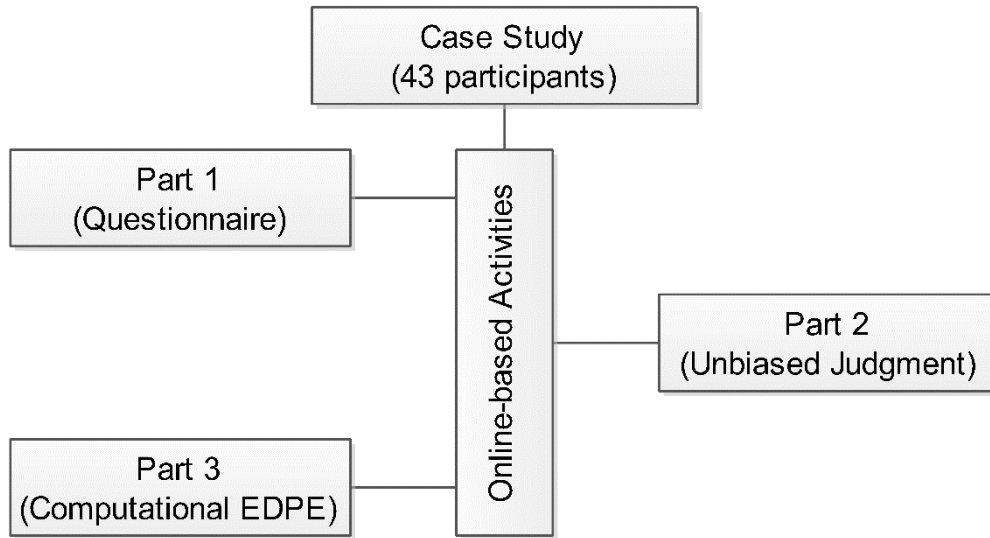
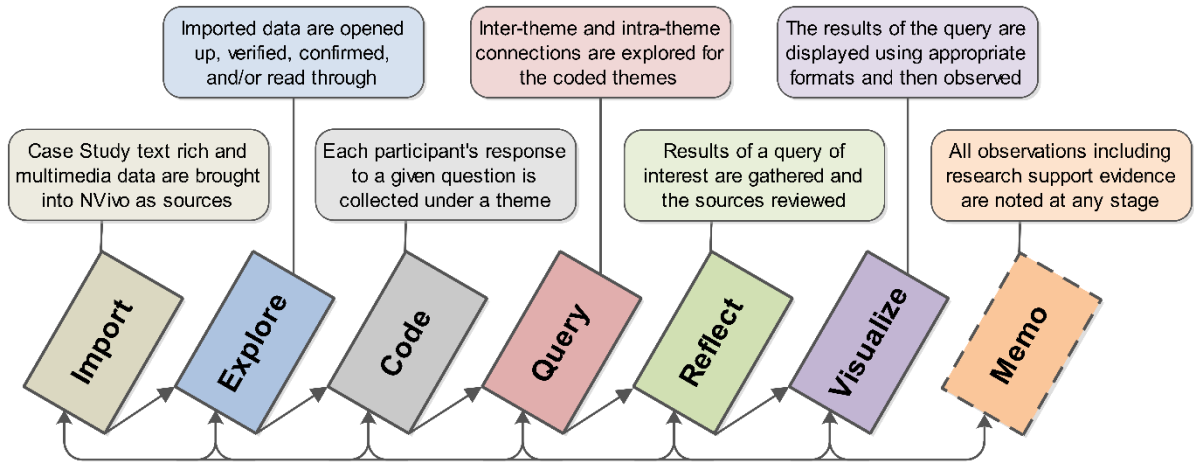


Fig. 5 Overview of data collection approach

1159

1160



1161

1162

Fig. 6 Case study qualitative analysis workflow

1163

1164

1165

1166

1167

1168

1169

1170

1171

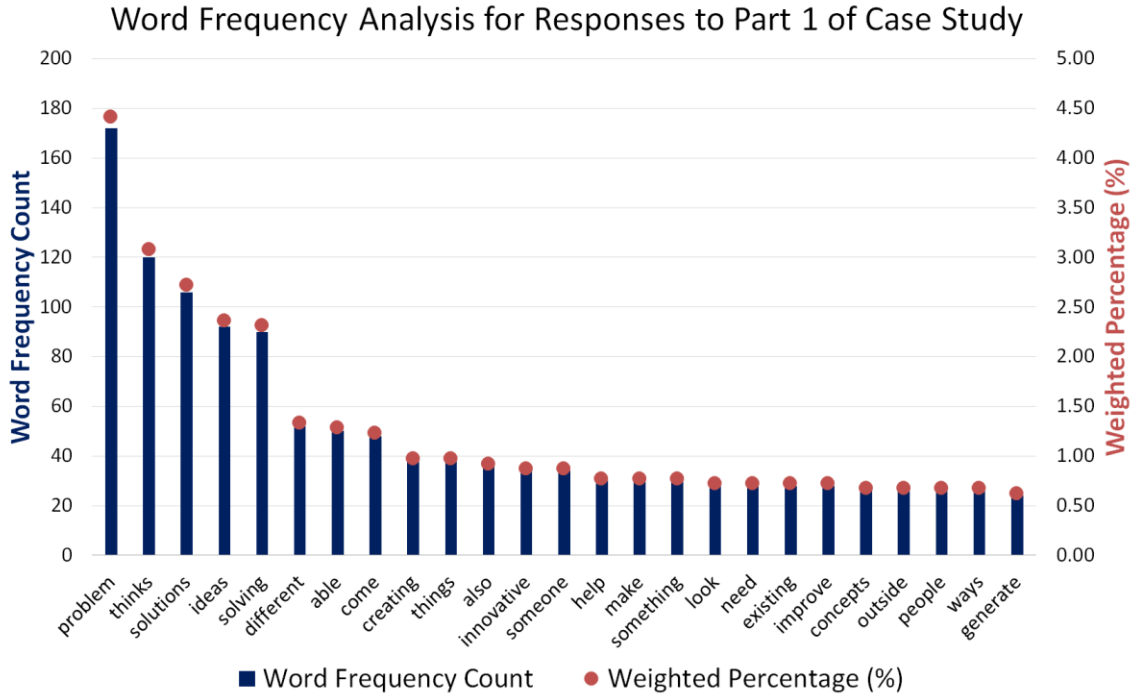
1172

1173

1174

1175

1176



1177

1178

Fig. 7 Most frequent words in explaining creativity in engineering design

1179

1180

1181

1182

1183

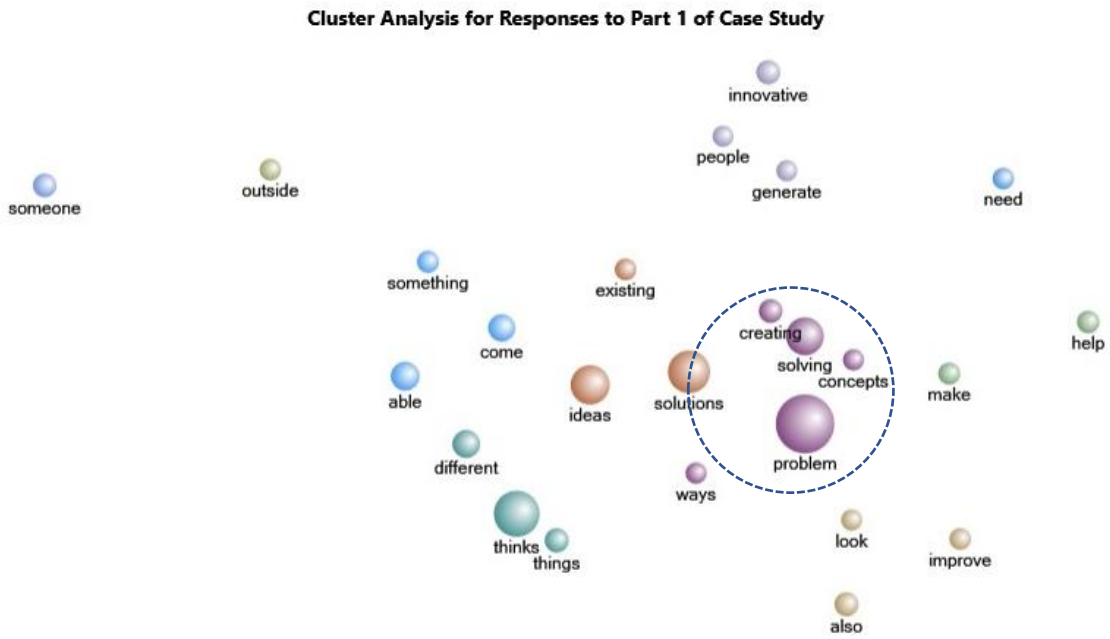
1184

1185

1186

1187

1188



1189

1190

Fig. 8 Word cluster showing a relationship in words explaining creativity

1191

1192

1193

1194

1195

1196

1197

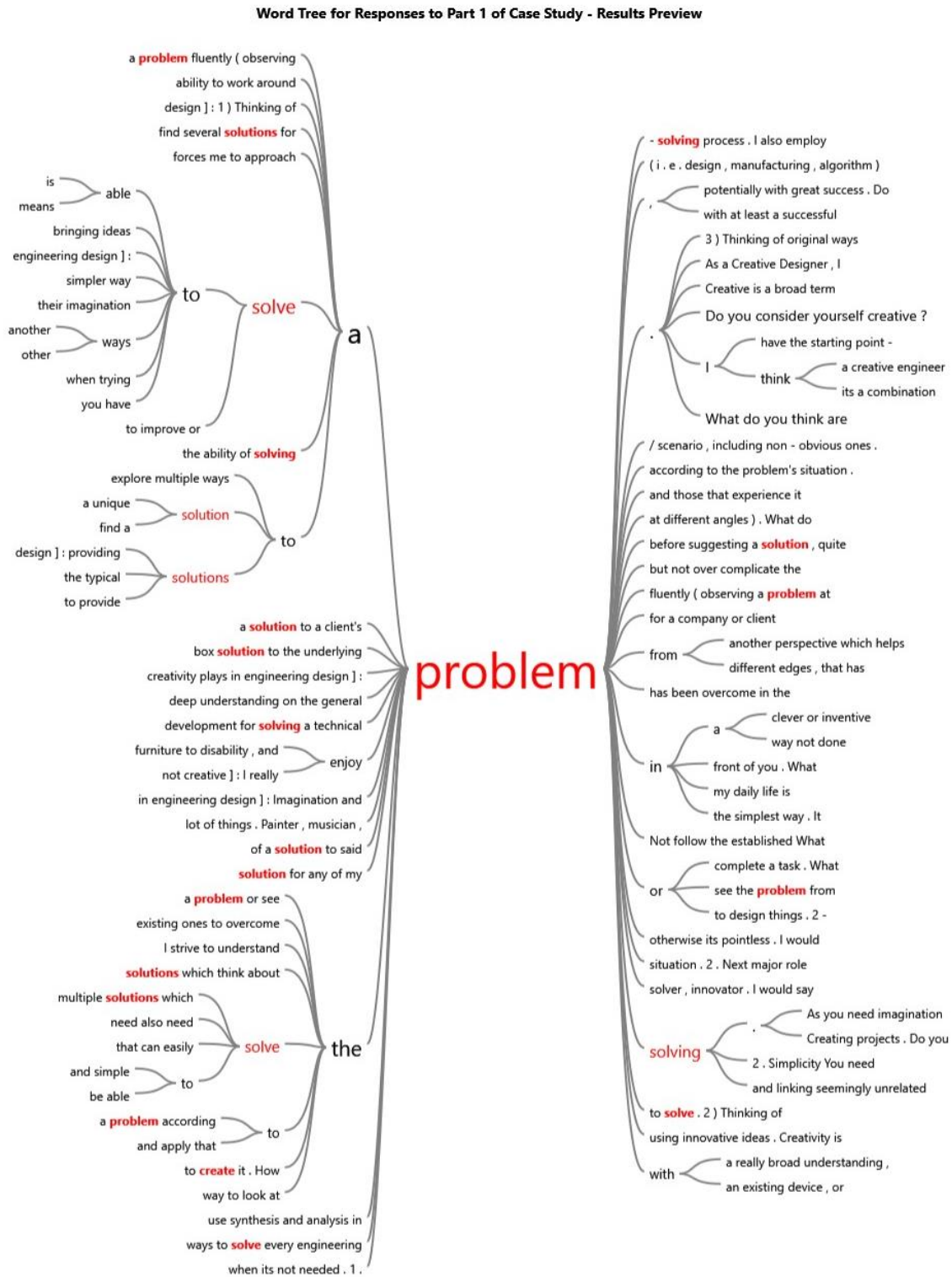
1198

1199

1200

1201

1202



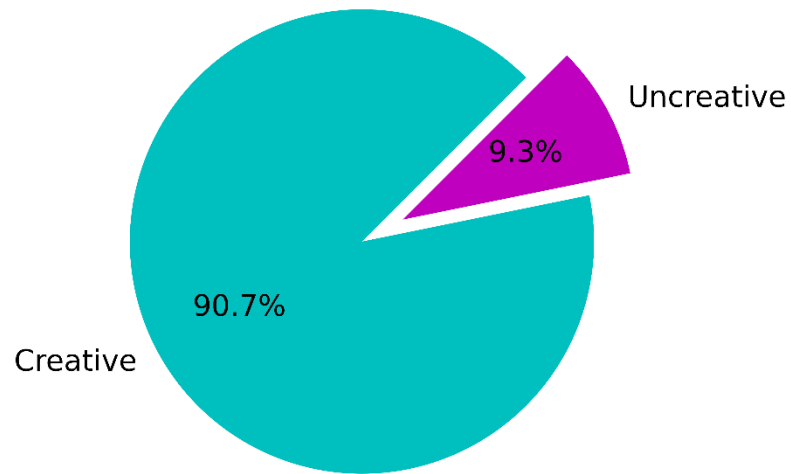
1203

1204

Fig. 9 Word Tree result for a text query search

1205

Participants Who Consider Themselves To Be Creative or Uncreative



1206

1207

Fig. 10 Participants' responses to their creativity ability

1208

1209

1210

1211

1212

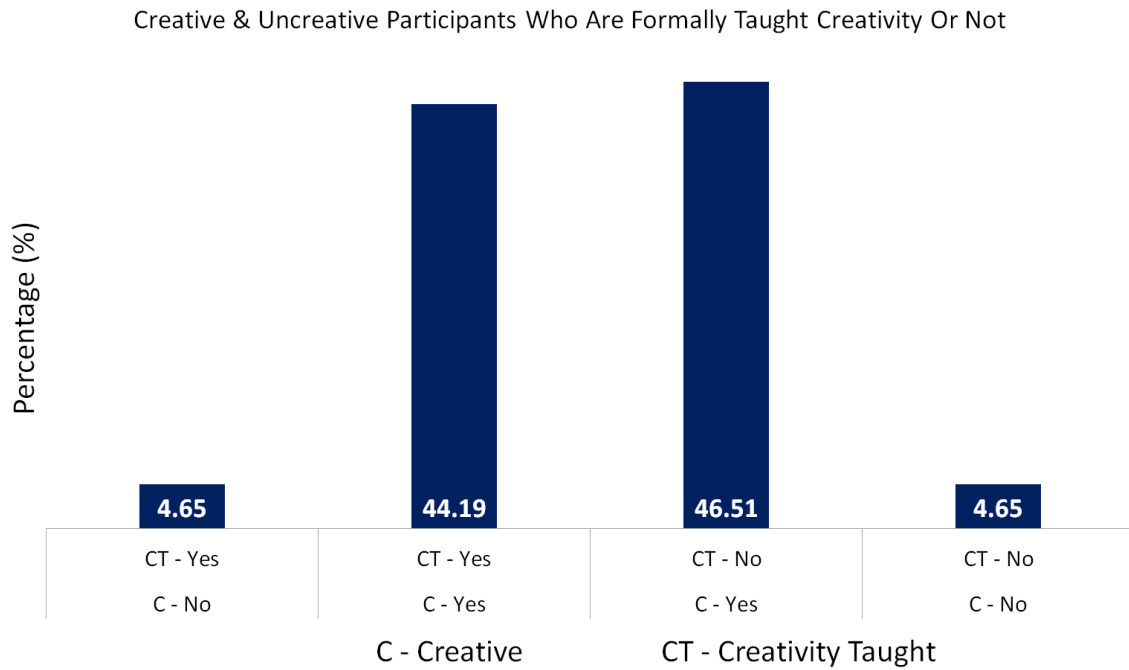
1213

1214

1215

1216

1217



1218

1219

1220

Fig. 11 An insight into creativity teaching in academia

1221

1222

1223

1224

1225

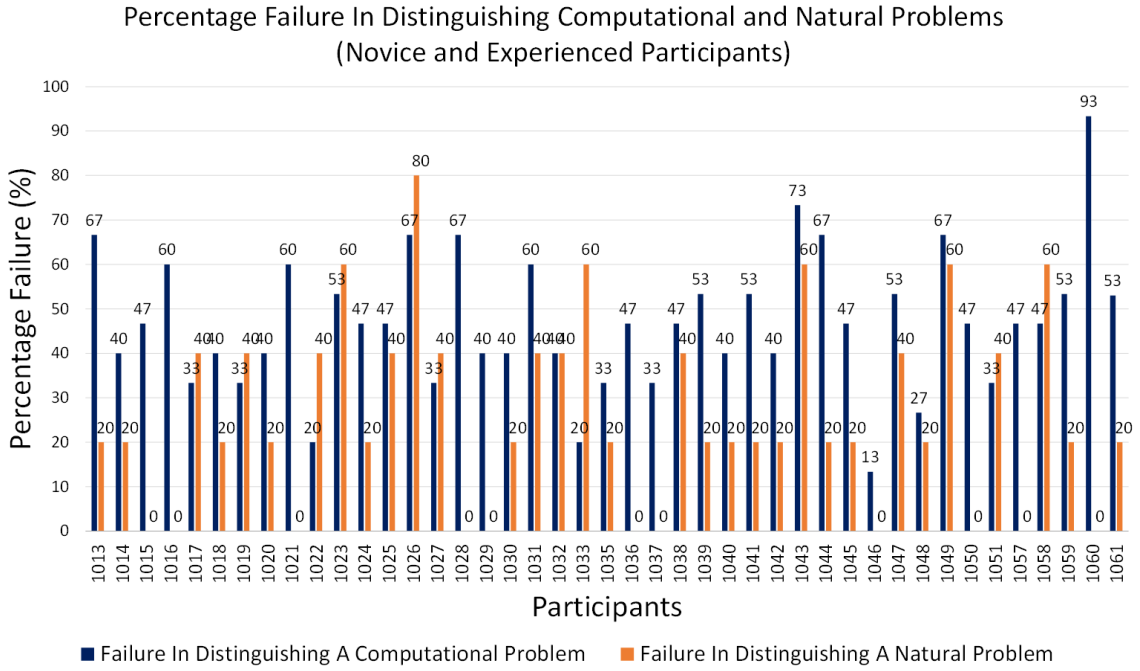
1226

1227

1228

1229

1230



1231

1232

Fig. 12 Result of Part 2 (Unbiased Judgement) of the case study

1233

1234

1235

1236

1237

1238

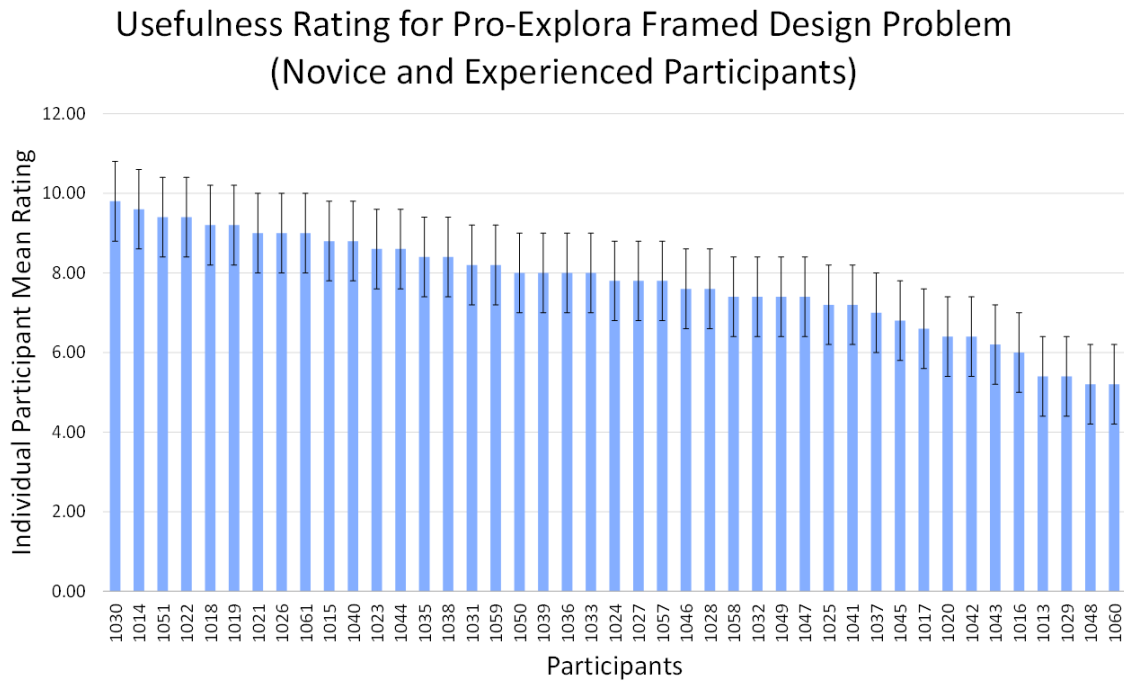
1239

1240

1241

1242

1243



1244

1245 **Fig. 13** Participants' usefulness rating for Pro-Explora framed EDP (with ± 1 SE bar)

1246

1247

1248

1249

1250

1251

1252

1253

1254

1255

1256

Table 1 Comparison of algorithms for natural and computational EDPE

Natural EDPE approach	Computational EDPE approach
Identify an EDP of societal relevance by accident (serendipity), stochastic synthesis (“garbage can” model), logical progression (“rational” model), and/or conceptualization (apophenia)	Frame an EDP of societal relevance by stochastic synthesis of big data, computational technologies (data extraction, ML, NLP), coding capabilities, connectionist theory, deterministic chaos, MM, BERT, and/or LSTM
Search manually for prior existence in relevant databases using search engines.	Make an automated search for prior existence in relevant databases using duplication recognitions.
Decide, subject to acceptance by the society or a relevant authority	Decide, subject to a design engineer's acceptance.

1257

1258

1259

1260

1261

1262

1263

1264

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277

1278

1279

1280

1281

1282

1283

Table 2 Case study participants' detail

Participants	Gender			Academic Qualification (Obtained/in view)			Experience (Years)	
	Male	Female	Total	B	M	PhD	≤ 3	> 3
Novice	12	7	19	5	11	3	19	0
Experienced	23	1	24	10	9	5	0	24
Total	35	8	43	15	20	8	19	24

1284

B – Bachelors, M – Masters

1285

1286

1287

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

1298

1299

1300

1301

1302

1303

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

1317

Table 3 Questionnaire for Part 1 of the case study

Questionnaire themes

- (a) What does it mean to be creative?
 - (b) Major are the roles of creativity in engineering design?
 - (c) Do You Consider Yourself Creative?
 - (d) Why are you creative or uncreative?
 - (e) Were you taught creativity at University or at work?
-

1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352

1353

Table 4 A sample set of 20 EDP for participants

Naturally and computationally framed EDP
1. Design of a mechanical intrusive force detection device.
2. To design a portable water distillation device.
3. A sustainable packaging design for wine.
4. Designing an interactive interface for collaborative engineering design.
5. A design of an automatic bottle opener.
6. Towards intelligent emotion detection system for video traffic surveillance.
7. Ai-based learning models for video traffic surveillance.
8. Design and material properties to minimize biofilm deposits.
9. Design of human-powered hybrid electric-power shovel for the physically challenged.
10. Design of self-reconfigurable production equipment during operation.
11. Anti riot drone without traffic lights.
12. Investigation of anomaly detection in a critical materials.
13. Design of a self-timing solar seawater desalination machine.
14. Staging co-design for reverse modeling of product development.
15. Detecting aggressive driving behavior using scilab.
16. Design of remote intelligent home finance software.
17. Designing products by artificial intelligence design approach.
18. A computationally efficient real-time vehicle and speed detection using federated learning.
19. Automatic mechanical footstep power tiller machine.
20. Design of production information retrieval system.

1354

1355

1356

1357

1358

1359

1360

1361

1362

1363

1364

1365

1366

1367

1368

1369

1370

1371

Table 5 Novice and Experienced participants

Novice				Experienced			
1013	1019	1026	1058	1022	1035	1041	1048
1014	1020	1027		1028	1036	1042	1049
1015	1021	1029		1030	1037	1044	1057
1016	1023	1043		1031	1038	1045	1059
1017	1024	1050		1032	1039	1046	1060
1018	1025	1051		1033	1040	1047	1061

1372

1373

1374

1375

1376

1377

1378

1379

1380

1381

1382

1383

1384

1385

1386

1387

1388

1389

1390

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404

1405

1406

1407

1408

Table 6 Correlation between experience and distinguishing a computational EDP

	P-value		Pearson's r	
	Novice	Experienced	Novice	Experienced
Experience vs Failed C differentiation	0.19	0.62	0.32	0.11
Experience vs Failed N differentiation	0.38	0.56	0.21	0.13

1409

C – computationally framed EDP N – naturally framed EDP

1410

1411

1412

1413

1414

1415

1416

1417

1418

1419

1420

1421

1422

1423

1424

1425

1426

1427

1428

1429

1430

1431

1432

1433

1434

1435

1436

1437

1438

1439

1440

1441

1442

1443

1444

1445
1446

Table 7 Cosine similarity assessment result

	EDP6	EDP7	EDP8	EDP9	EDP10	Q*
EDP1	0.8163	0.7000	0.7461	0.8465	0.8589	0.6163
EDP2	0.7164	0.6240	0.7838	0.6794	0.6686	0.5628
EDP3	0.6700	0.6656	0.6258	0.6776	0.6872	0.5249
EDP4	0.7842	0.7275	0.7191	0.7500	0.7710	0.5371
EDP5	0.6163	0.5767	0.5481	0.7164	0.7306	0.4750

1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478

1479 **Table 8 Correlation between experience and rating of a computational EDP**

	P-value	Pearson's r
Experience vs Usefulness Rating (Novice)	0.16	-0.34
Experience vs Usefulness Rating (Experienced)	0.69	-0.08

1480