Studying the Influence of Empathy Maps on Brainstorming for Requirements Elicitation: A Quasi-Experiment

Ezequiel Kahan
Universidad Nacional de 3 de Febrero
Buenos aires, Argentina *ekahan@untref.edu.ar*

Emilio Insfran
IUMTI - Universitat Politècnica de València
Valencia, Spain *einsfran@dsic.upv.es*

Marcela Genero
University of Castilla-La Mancha
Ciudad Real, Spain *marcela.genero@uclm.es*

Alejandro Oliveros
Universidad Nacional de 3 de Febrero
Buenos Aires, Argentina *aoliveros@untref.com.ar*

Abstract

There is currently a trend that highlights the interest in the use of Design Thinking techniques for the elicitation of requirements. The evidence concerning the effectiveness of these techniques is, however, still scarce, which has consequently led us to carry out a quasi-experiment in order to evaluate the effectiveness of one of the most frequently used DT techniques: Empathy Maps. Empathy Maps were employed by students enrolled on a Bachelor's degree in Computer Engineering at the Universitat Politècnica de València in Spain in order to understand their influence on the effectiveness of Brainstorming sessions, which was measured in terms of the quantity of requirements attained. The students perceived a slightly positive influence as regards the use of the Empathy Maps as part of a requirements elicitation, despite not finding statistically positive results in favour of Empathy Maps. We obtained several insights that may be useful for practitioners, researchers and lecturers interested in using this type of techniques to improve the requirements elicitation practices currently employed to develop information systems, although this preliminary result is not conclusive and should be corroborated in further studies.

Keywords: Requirement elicitation, Design thinking, Empathy Map, Experiment.

1. Introduction

The first work that linked Design Thinking (DT) to Requirements Engineering (RE) appeared more than 8 years ago [27]. The interest in this topic has since increased, and there is now a solid body of studies demonstrating the potential of applying DT in synergy with RE in general, and to the requirements elicitation phase in particular [8], [11], [14, 15]. DT is considered to be a "*way of finding human needs and creating new solutions using the tools and mindsets of design practitioners*" [16]. DT can be integrated into business and other activities of society, and individuals and teams can use it to generate innovative and implementable ideas that, that have an impact [3]. It is consistent with the initial elicitation goals and practices of RE and the rapid prototyping and customer involvement of agile development methods [27]. The DT process, which is complemented by design tools and techniques, can provide a supportive framework for information system (IS) development by enabling teams to invent new constructs and artefacts [12].

In the field of IS, research efforts related to DT take two principal directions: (1) how IT and IS can support DT practices by generating new tools, and (2) how DT can support problem-solving in the IS context [12]. With regard to this second direction, DT has the potential to solve RE challenges [13], [20], including the discovery of the fuzzy needs and volatile

requirements of the multiple stakeholders involved in socio-technical system projects [10]. Evidence shows that it would be possible to improve the application of DT in RE [12], and this can specifically be seen in the elicitation of requirements to which a fragmented perspective of DT is applied, and in which the emphasis is more on techniques than on its integrative perspective. Furthermore, the evidence that exists on its specific techniques is generic, that is, the empirical evidence regarding which of the techniques that are part of the DT process are more appropriate or have better results when the aforementioned problems are confronted. This provides the possibility of conducting experiments with which to evaluate the relevance of a process based on DT in order to incorporate it into requirement elicitation, and the different techniques that form part of it.

This paper, therefore, presents the results of a quasi-experiment carried out in order to investigate whether the use of Empathy Maps in combination with Personas improves the performance of Brainstorming sessions. The improvement was measured in terms of the quantity of requirements obtained in the ideation sessions. We decided to begin by carrying out an evaluation in an academic context. As indicated in [5], experiments with student participants must be considered state of the art for evaluation in RE as a prior and necessary step before carrying out evaluations in industry. The quasi-experiment was conducted with a group of 72 students enrolled on a Bachelor's degree in Computer Engineering at the Universitat Politècnica de València in Spain.

The remaining of this paper is structured as follows. Section 2 introduces a process that incorporates DT techniques for requirements elicitation, while Section 3 provides an overview of the related work. Section 4 introduces a definition of the quasi-experiment, along with its planning and execution while Section 5 presents the analysis and interpretation of the data collected in the quasi-experiment, and Section 6 discusses the threats to validity. Finally, Section 7 presents our conclusions and outlines our future work.

2. Requirements elicitation process based on Design Thinking

We propose to use our previous experience [14] as a basis on which to incorporate a process based on the Design Thinking Stanford approach (empathize, define, ideate, prototype, test), which is one of the approaches most frequently used by practitioners [21], into the requirements elicitation phase. Our proposal is based on that of Hehn et al. [9], who introduce multiple strategies for the integration of DT into RE. Of the 5 stages in this approach, we consider the first 3, since we consider that they are the most closely connected with the requirements elicitation phase [15] (see Figure1).

We consider that *empathy* is a key feature of requirements elicitation and also of DT [23]. Empathy is a concept that includes both the involuntary act of feeling sympathy for someone else and the cognitive act of placing oneself in someone else's position and adopting their perspective [17]. It is the attempt to reconstruct the specific perspective of the other and how he or she perceives the situation. For example, employing empathy techniques can foster software engineers' empathy practices and skills, and help identify requirements that are strongly focused on people, such as privacy requirements [18]. It can be viewed as a form of knowledge construction [17], which is a central aspect of requirements elicitation [19].

Although empathy development occurs in all stages of DT, it is particularly relevant during the "Empathise" stage, which is usually the first stage of the process. Nonetheless, empathy occurs in the entire requirements elicitation process.

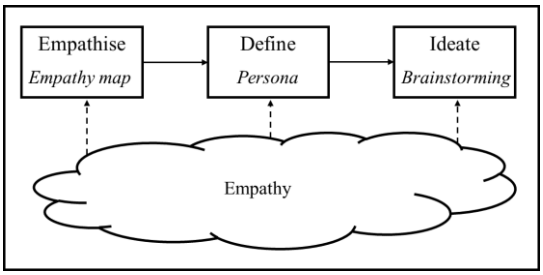


Figure 1. A process for requirement elicitation based on DT

As shown in Figure 1, we propose to use a specific technique in each stage: an Empathy Map, Personas and Brainstorming, respectively. According to the existing literature [15], there

are at least 70 DT techniques, although several of them are similar to each other, and it is common to use a small group of them [15]. Our initial view is that of empathising using an Empathy map, after which synthesising / defining information using Personas could improve the ideation process, which is carried out by means of Brainstorming. These techniques were chosen on the basis of an exhaustive review of the related literature [14, 15], [25] and the experience of the first author, who has been applying DT as a practitioner in the software industry in Argentina since 2012. We consider that the incorporation of the Empathise stage is a key aspect of an elicitation process based on DT, which is not necessarily present in the more classical requirements elicitation process, or when DT techniques are used individually. Although other authors have referenced several DT techniques, such as Personas and Customer Journey Maps, as effective means for requirements elicitation, they have not addressed the users' emotional and cognitive aspects, but rather the functional requirements and goal aspects of the system [18]. These emotional and cognitive aspects could, in our opinion, be in part achieved by incorporating techniques frequently used during the empathy stage, such as Empathy Maps. Despite the fact that the template for Personas development has, in some cases, included Empathy map attributes [7], these elements, which are a fundamental part of empathy, have not been covered in any depth.

The requirement elicitation process based on DT illustrated in Figure 1 consists of the following stages and techniques:

- Empathise: During this stage, the goal is to explore the problem space and attempt to understand it by carrying out different activities that include research, consulting domain experts, observing, engaging, and empathising with people. Some of the techniques most frequently employed during this stage are interviews, focus groups, observations, documentary analysis and Empathy Maps. The technique selected for this stage was that of an Empathy Map, since it would make it possible to integrate learning and generate valuable "insights" about the user. An Empathy Map is a simple and easy-to-understand visual representation that allows the knowledge regarding the behaviour and attitudes of one or a group of users to be captured, thus generating empathy. The exercise of creating the map helps participants to consider things from the user's perspective, along with his/her goals and challenges. The map provides four central areas on which to focus: what the user (s) said, did, thought or felt in relation to the problem to be solved.
- Define: During the define stage, the information gathered during the empathise stage converges on the definition of a specific problem. This stage is also denominated as synthesis, because it involves reframing the problem for a specific group of users, environment, or situation. The technique selected for this stage was that of Personas. Personas are fictional characters, which are created on the basis of previous research in order to represent the different types of users that might use a solution. Creating personas helps to understand one's users' needs, experiences, behaviour and goals. The Personas method is used in the design of digital environments, but is also employed in many other contexts, including product development, marketing, communication planning and service design.
- Ideate: This is the stage in which solutions are generated and proposed, in our case, in the form of ideas of requirements. The technique selected for this stage was that of Brainstorming. Brainstorming is a group creativity technique in which solutions to a specific problem are proposed by gathering a list of ideas spontaneously contributed by the group's members. Inhibitions as regards participation are removed by encouraging people to think more freely and to suggest as many spontaneous new ideas as possible without criticism. Note that we prefer to use the term *idea of requirements* rather than *requirements* when referring to the result of the Brainstorming technique, since it represents raw ideas of requirements that still need to be introduced to and validated with the stakeholders if they are to be considered as the requirements of the information system to be developed.

In addition to these stages and specific techniques, we also consider that DT, as an holistic approach, provides additional enrichment based on empathy building, focusing on user needs and the frequent validation of ideas through the use of prototypes [27].

3. Related work

The interest in employing DT techniques for requirements elicitation has grown in the IS field in recent years[8, 15]. We, therefore, used the empirical studies found in the two aforementioned SMSs, which cover papers published until 2018, as a basis for our research. Scopus was subsequently used to extend this search until December 2020, which enabled us to discover 5 more empirical studies published from 2017 onwards that show evidence of the use of specific DT techniques in requirements elicitation. All of the studies found are described below. Three of the studies were carried out in an academic context at Brazilian universities [4, 7, 26]:

- Costa et al.[4] carried out an exploratory case study with 17 undergraduate Computer Science students with the aim of discovering: "What are the perceptions of students regarding learning Design Thinking?". Projects using individual techniques (Personas, Empathy Maps) and team techniques (Brainstorming and co-creation workshop) were then employed for the development of the authors' mobile application. The students considered that the techniques used were very useful but stated that more training time was required to carry out the case study.
- Souza et al. [26] presented exploratory case studies performed with 32 and 21 undergraduate students, respectively. The authors selected 15 techniques with which to teach their students that they should elicit the requirements of a specific system of their own choosing. The techniques selected were grouped according to their purpose function: 1) Collection, registration, and organisation of information: exploratory research, rapid ethnography, fly on the wall, interviews, questionnaires, insight cards, Empathy Map, Personas and motivation matrix; 2) Group use: Brainstorming and bodystorming; 3) Stimulating the generation of ideas and representing possible solutions: storytelling, storyboards, prototyping and try it yourself. In both experiences, the students reduced the number of techniques used: 6 in the first and 12 in the second, with Personas and Brainstorming being those most frequently used, followed by Empathy Maps.
- Ferreira et al. [7] presents a controlled experiment carried out with 37 Computer Science undergraduates in order to compare two Personas-related techniques: traditional Personas and PATHY, which is the traditional Personas technique enriched with some concepts of Empathy Maps. The controlled experiments assessed which of the two techniques helped to generate descriptions of personas that are more focused on potential requirements to be considered in the design of the application. The authors analysed the efficiency of the techniques and the participants' perceptions of their use. PATHY generated more relevant characteristics for the application design than did the technique that followed the traditional description. PATHY was also more efficient as regards creating personas.

Three studies were performed in an industrial context with practitioners [10], [22], [24]:

- Hehn and Uebernickel [10] performed an exploratory case study, carried out by interviewing 18 practitioners. Their findings suggest that DT has the potential to support current RE practices and vice versa. The authors in question believe that an integration of both approaches leverages a symbiotic relationship. As both topics, DT and RE, are broad-ranging, multiple opportunities for further research can be identified. DT provides a structured process for requirements elicitation for complex problems.
- Ruchira et al. [24] performed a case study within the development of agile software projects in the Sri Lankan Software Development Industry. After carrying out 15 interviews with project stakeholders who practice DT in agile projects, they identified that DT practices, such as the human-centred approach, thinking by doing, visualising, the synthesis of diverging and converging, and a collaborative work style, are helpful as regards enhancing the customer's expectations. The most suitable methods by which to identify customers' needs were determined to be: customer journey, story mapping, prototypes, proof of concept and customer profiling.
- Prestes et al. [22] carried out a survey of 127 practitioners on the use of DT in agile software development in Brazil, with the aim of answering the following research question: "How to apply DT practices to improve customer expectations in agile practices?". The results show that a great variety of techniques (more than 50) and different variants of the DT process are

employed, with Brainstorming, Personas and Empathy Maps being those most frequently used.

As stated above, the number of publications showing some type of evidence of the use of DT in requirements elicitation has increased in recent years, particularly in Brazilian institutions. Studies conducted in an academic context show that the DT techniques for which there is the most empirical evidence of their use are Personas, Brainstorming and Empathy Maps. The studies carried out in an industrial context are, however, more generic, i.e., they do not focus on specific techniques, although the evidence is not very rigorous, since controlled experiments are not carried out or different techniques are not compared but simply used and some type of subjective feedback based on surveys or interviews is collected. It is, therefore, necessary to collect evidence regarding which specific techniques are the most appropriate for each of the stages related to requirement elicitation within the DT requirement elicitation process. This led us to propose a process based on DT as described in section 2 (see Figure 1), together with specific techniques for each of its phases and to validate it empirically.

4. Quasi-Experiment description

The quasi-experiment was carried out at the Universitat Politècnica de València (Spain) in October 2019. A quasi-experiment was carried out rather than a "true" experiment because the random assignment of the treatments to the participants was not possible. This quasi-experiment was designed and reported by following the recommendations provided in [28]. The experimental material is available at: https://tinyurl.com/js4m2xj7, and the principal characteristics of the quasi-experiment are described in the following subsections.

4.1. Goal

The main goal of this quasi-experiment was to investigate whether a sequence of *elicitation techniques* influences the effectiveness of subjects when eliciting requirements. The Goal Question Metrics template [1] was employed, and the goal of this quasi-experiment was, therefore, **to analyse** Elicitation Techniques **for the purpose of** comparing them **with respect to** their effectiveness, perceived ease of use, and perceived utility **from the point of view** of requirements analysts **in the context** of students enrolled on a Bachelor's degree in Computer Engineering at the Universitat Politècnica de València in Spain.

4.2. Variables and hypotheses

The independent variable of interest is the *elicitation technique* used (Empathy Map + Personas + Brainstroming or Personas + Brainstorming); i.e., two treatments were considered (*EM+P+B and P+B*).

The dependent variable was *Effectiveness*, which was measured as being the quantity of ideas of requirements (QIR) generated by the students. These QIR were grouped into the following categories:

- People-centred ideas. This category included all the ideas that impact on or are impacted on by people. They include ideas in which people are the central component. For example: "The application will be accessible to all users, including those with functional diversity".
- Business-oriented ideas. This category included all the ideas that describe or propose aspects related to an animal protection centre (the domain chosen for use in this quasi-experiment, as explained below). For example: "Include an inventory that shows the food and medicines available for both purchase and donation".
- Technology-oriented ideas: This category included all the ideas that establish technological requirements. For example, "The public part of the application must be a website".
- Others: Ideas of requirements that do not connect directly with the software application itself. For example, "Creation of tutorials on how to properly care for animals".

The sum of the number of ideas of requirements in each category was, therefore, considered as a measure of *Effectiveness* (QIR = QIR-P + QIR-B + QIR-T + QIR-O). We formulated the following hypotheses:

- H1-0: There is no significant difference between the subjects' effectiveness when using *P+EM+B or P+B* / H1-a: ¬ H1-0.

The goal of the statistical analysis is to reject the null hypothesis and possibly accept the alternative ones. The hypothesis presented herein is two-sided, because we did not postulate any effect arising.

4.3. Sample and participants

We took a convenience sample of two groups of undergraduate students enrolled on a Bachelor's degree in Computer Engineering at the Universitat Politècnica de València (UPV) in Spain (Group A composed of 35 participants and Group B composed of 37 participants). The students attended a course on Requirements Engineering during the academic year 2019-2020. This course included an introduction to and examples of the use of the techniques employed in the quasi-experiment, i.e. Personas, Empathy Maps and Brainstorming. The participants had no prior experience in the use of any of these three techniques. Finally, considering that the main purpose of the quasi-experiment was to study the improvement to the quantity of ideas of requirements generated during the Brainstorming session when using Empathy Maps, and that Brainstorming is a group-based technique, we defined several working groups with which to run the quasi-experiment. Each of the working groups, almost all of which were composed of 5 students, was randomly assigned by the course professor. Only two working groups in Group B had 6 students owing to the number of students enrolled on the course, signifying that there were 7 working groups in Group A (35 students) and 7 working groups in Group B (37 students).

4.4. Experimental objects and tasks

The object of the experiment was to describe the characteristics and main needs of an Animal Adoption Centre, called MODEPRAN, which is in the city of Valencia in Spain. This description provided the participants with an overview of and context in which to begin identifying the main stakeholders, and the scope in which to propose the ideas for the software system requirements during the Brainstorming session. This domain was chosen because it was possible that most of the participants would be familiar with the process of adopting animals, which might motivate and inspire interaction among them.

The experimental task included the generation of ideas of requirements by means of a Brainstorming session using only Personas, in the case of Group B **(**P+B treatment**)**, or using the Personas together with the Empathy Maps, in the case of Group A **(***EM+P+B treatment***)**. The rationale behind this task configuration was that using Personas together with the Empathy Maps might improve empathy with the stakeholders and consequently improve the quantity of ideas of requirements generated during the Brainstorming session. Multiple documents were defined as instrumentation. The documentation included: i) a description of the MODEPRAN Animal Adoption Centre; ii) an introduction to the Personas technique with examples; iii) an introduction to the Empathy Maps technique with examples; iv) guidelines on how to run a Brainstorming session and the format required in order to report the ideas of requirements generated; and v) the post-experiment survey, which included both closed questions in order to analyse the perception-based variables of this study (PEOU, PU and MT) and one open question that would enable the participants to express their opinion about the use of the techniques. Table 1 shows the closed questions in the post-experiment survey. The questionnaire contained three questions related to the students' perceptions of the material and training they had received, while the remaining questions were related to the students' perceptions when performing the experimental tasks using each treatment. These questions were based on an adaptation of the Technology Acceptance Model (TAM) [6], and measured the Perceived ease of use (PEOU) and Perceived utility (PU) using a five-point Likert scale. The number of questions related to PEOU and PU were different for each group because Group A used the Empathy Map technique, and specific questions related to this technique were, therefore, added to the questionnaire.

Table 1. Post-experiment questionnaire

ID	Classification	Group A	Group B
Q1	Material and training (MT1)	The support material used in the exercise was adequate to be able to understand and use each of the techniques (Personas, Empathy Map, Brainstorming).	The support material used in the exercise was adequate to be able to understand and use each of the techniques (Personas, Brainstorming).
Q2	Material and training (MT2)	The explanation received was adequate in order to apply each of the techniques (Personas, Empathy Map, Brainstorming).	The explanation received was adequate in order to apply each of the techniques (Personas, Brainstorming).
Q3	Material and training (MT3)	I now have the knowledge required to be able to use the techniques learned (Personas, Empathy Map, Brainstorming) in real situations.	I now have the knowledge required to be able to use the techniques learned (Personas, Brainstorming) in real situations.
Q4	Perceived utility (PU1)	The Personas technique was useful for characterising the different stakeholders.	The Personas technique was useful for characterising the different stakeholders.
Q5	Perceived utility(PU2)	The Brainstorming technique was useful for identifying requirements ideas.	The Brainstorming technique was useful for identifying requirements ideas.
Q6	Perceived ease of use (PEOU1)	There was a good connection between the different techniques used (Personas, Empathy Map, Brainstorming).	There was a good connection between the different techniques used (Personas, Brainstorming).
Q7	Perceived ease of use (PEOU2)	The sequence in which the techniques were applied (Personas, Empathy Map, Brainstorming) was appropriate.	The sequence in which the techniques were applied (Personas, Brainstorming) was appropriate.
Q8	Perceived ease of use (PEOU3)	It was easy to construct the Empathy Map.	n/a
Q9	Perceived ease of use (PEOU4)	The construction of the Empathy Map did not require a great effort.	n/a
Q10	Perceived ease of use (PEOU)	It was easy to separate the information into the 4 quadrants of the Empathy Map.	n/a
Q11	Perceived utility (PU3)	The Empathy Map technique was useful for exploring stakeholder goals and motivations.	n/a
Q12	Perceived utility (PU4)	Constructing the Empathy Map was beneficial for the Brainstorming session.	n/a

4.5. Design and execution

We used a between-subject design, meaning that the participants (i.e., working groups) in the quasi-experiment were assigned to different treatments, with each working group experiencing only one of the treatments. The students were not aware that they were participating in a quasi-experiment and were under the impression that it was just another exercise in the context of the RE course on which they were enrolled.

Since the RE course is a weekly course of three hours per week, the training and quasi-experiment were performed in two sessions throughout two weeks. The first week was the training session, whose purpose was to introduce the concepts, examples and small exercises concerning the techniques that would be applied in the quasi-experiment: Personas, Empathy Maps, and Brainstorming in Group A and Personas and Brainstorming in Group B. The quasi-experiment took place in the second week.

The quasi-experiment was controlled, meaning that no interactions took place between the working groups. The experimental session lasted approximately three hours. In treatment A, the first 90 minutes were dedicated to defining the Personas and the Empathy Maps, while the last 90 minutes were dedicated to the Brainstorming session, including the generation of the ideas of requirements and the consolidation and prioritisation of the results. Since the groups involved in treatment B did not use the Empathy Maps, only 60 minutes were dedicated to defining the Personas, while the remaining time (120 minutes) was employed to run the Brainstorming session. Once the quasi-experiment had finished, in order to prepare the data for the analysis, two of the authors of this paper classified the ideas of requirements obtained by each of the working groups according to the classification introduced in Section 4.2 (i.e. people-centred, business-oriented, technology-oriented, others). Each idea of requirements was analysed and classified into one or more category, with the authors reaching a consensus when necessary.

4.6. Analysis procedure

The data analysis was carried out by considering two phases:

1) Analysis of Effectiveness measured through number of QIR attained

- We first carried out a descriptive study of the measure of the dependent variable, i.e. QIR, in order to obtain a general overview of the data obtained in the quasi-experiment.

- We performed a Kolmogorov-Smirnov test [29] to determine the normality of distributions.
- The results of the afore mentioned test were employed as a basis on which to test the hypothesis formulated, using the non-parametric Mann Whitney test [37] for the data collected in the quasi-experiment.

2) Analysis of the post-experiment survey

The perception-based measures collected from this survey were analysed using descriptive statistics illustrated by box-plots. The analysis was grouped into PEOU, PU and MT questions.

5. Data Analysis and Interpretation

In this section, we present the data analysis and interpretation of the results obtained in the quasi-experiment following the analysis procedure detailed above.

5.1. Analysis of the Effectiveness

Table 2 reports the descriptive statistics of the ideas of requirements classified.

Table 2. Descriptive statistics of the ideas of requirements classified.

	Treatment EM+P+B (Group A)					***Treatment P+B (Group B)***				
	Media	*Median*	*SD*	*Min*	*Max*	*Media*	*Median*	*SD*	*Min*	*Max*
QIR	30.86	28	10.792	20	47	31.43	28	11.013	18	48
QIR- P	4.43	4	3.409	1	11	4.43	4	2.37	1	8
QIR-B	22.43	18	8.182	15	36	21	20	9.661	11	36
QIR-T	4	4	2.38	2	9	6	6	3.416	1	11

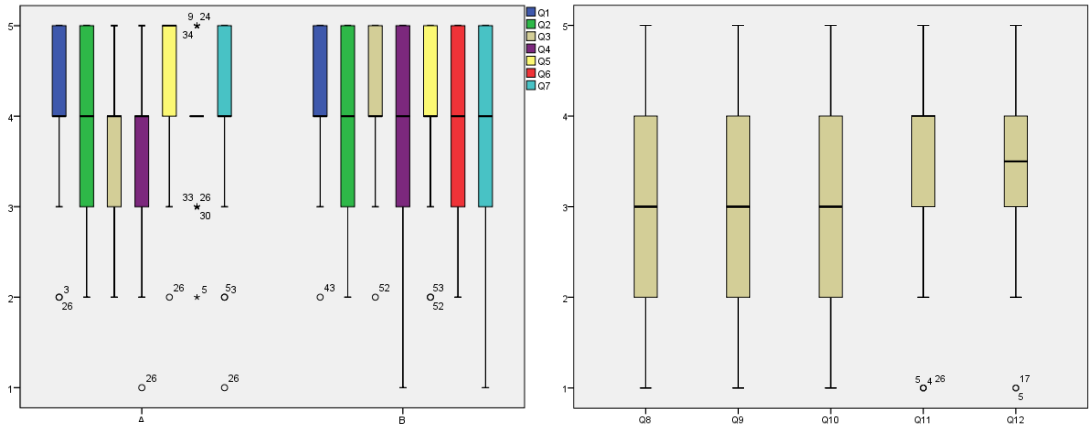
The descriptive statistics show that fewer technology-orientated ideas were generated by the groups in the Empathy Map treatment. One possible explanation for this is that the use of an Empathy Map influences the number of ideas linked to non-functional requirements, thus reducing their number. In order to test the hypothesis formulated, we analysed the effect of the main factor (Empathy Map) on the measures considered (QIR, QIR-P, QIR-B, QIR-T) using the non- parametric Mann-Whitney test. Table 3 shows the results obtained for each measure employed in the Mann-Whitney U tests, in which the Origin column describes the independent variable, p-value is the statistical significance obtained, op is the estimated observed power of the test, es is the effect size, and r represents the possibility of rejecting the null hypothesis with the data obtained. All these values were calculated using a standard configuration of SPSS. The results obtained do not allow us to reject H1-0: given that the p-value is 0.798, which is greater than 0.05., i.e. it would appear that the Empathy Map had no effect on the QIR. Similar results were obtained after repeating the test for each individual variable (QIR-P, QIR-B and QIR-T), i.e. it was not possible to reject H1-0 in any of the cases.

Table 3. Mann-Whitney test results for QIR

	Requirement elicitation technique			
	p-value	**op**	**es**	**R**
QIR	.798	<0.051	-0.098	NO
QIF-P	.846	<0.050	0	NO
QIF-N	.608	<0.059	0.299	NO
QIF-T	.221	<0.216	-1.271	NO

5.2. Analysis of the post-experiment survey

Figures 2 and 3 show the box-plots generated with data obtained from the post-experiment survey. The contents of these figures are analysed separately for each question as follows.



A) Questions 1 to 7

B) Questions 8 to 12 (applied to Group A only)

Figure 2. Box-plots of questions of the post experiment-survey

- Q1 - Upon evaluating the answers, it would appear that the material used is perceived as adequate, although the use of the Empathy Map generates a slightly lower average value.
- Q2 - Upon evaluating the second answer, it follows that the explanation of the techniques provided by the teacher was adequate. There is more dispersion in the answers obtained for this question than in the case of Q1. A possible explanation for this greater dispersion could be the difference in the students' preferences as regards the learning modality. The evaluation of Q1 and Q2 suggests that the teacher's explanation is consistent, since there were no differences between the explanations given to both groups.
- Q3 - The perception of the level of knowledge is lower in the Empathy Map treatment. This suggests that the Empathy Map, by requiring more dedication and effort (see specific questions in the next section), allowed the students to attain a more in-depth idea about their knowledge (they were more aware that they did not know as much).
- Q4 - The students perceived the Personas technique to be less useful in the Empathy Map treatment. This suggests a certain level of redundancy between both techniques, signifying that the Persona technique was perceived as less useful in the Empathy Map treatment than in that in which it was not employed. It might also suggest that it is not necessary to apply both techniques. In order to discover more about this aspect, in the future, a third treatment could be incorporated in which only the Empathy Map and the Brainstorming session are carried out so as to compare whether there is a variation in the results.
- Q5 - The perception of the utility of the use of Brainstorming attained a higher value in the EM+P+B treatment, with a similar distribution of percentiles for the 2 treatments. Upon comparing the mean values attained by the two groups, it will be noted that the students who used the Empathy Map had a better perception of Brainstorming.

Table 3. Mean value of the perceived utility of the different techniques per Group/treatment

Technique	Group A	Group B
Brainstorming	4.55	4.18
Empathy Map	3.36	-
Personas	3.67	3.96

Although the PU of the Empathy Map is the lowest of the 3 techniques (see Figures 2 and 3), it would appear that its use improves the PU of Brainstorming. This suggests that the Empathy Map could help to order the Brainstorming session by providing a prior "qualitative" analysis of the stakeholders.

- Q6 - A comparison of both conditions shows that there is a lower standard deviation for the perception of connection between the techniques when the Empathy Map treatment is employed. This is an indication of the advantage of using the Empathy Map. It could also denote a better relationship between the student and the technique than in the other

case. A more stable result can be expected with its use, which contributes to its teaching in the educational field.

The evaluation of Q5 and Q6 suggests that "when the product is good, the process that leads to it is also good". Since the PU of Brainstorming in treatment *EM+P+B* is higher, the perception of the process as a whole is also higher, despite the fact that the process is more extensive and requires more effort in the *EM+P+B* treatment, and that the PU of the Empathy Map and Personas techniques attained a lower PU value than that of Brainstorming.

- Q7 - Upon comparing both treatments, we discovered that those students who used the Empathy Map perceived the sequences in which the techniques were used in the Empathy Map treatment to be the same, but that a lower standard deviation was attained for the *EM+P+B* treatment. Again, a more stable result could contribute to its teaching in the educational field.

The results obtained for the specific questions regarding the *EM+P+B* treatment are detailed below:

- Q11. The perception of utility of the empathy technique is lower than that of Personas and Brainstorming techniques. This is consistent with questions Q9, Q10 and Q11.
- Q8, Q9 and Q11. The median value of the PEOU for the specific questions concerning the EM+P+B treatment show that lower values were attained than for the PEOU concerning People and Brainstorming, which is consistent with a lower PU. There is much more dispersion when compared to the previous answers. One hypothesis is that, as the technique was difficult, it required more effort, and it was complicated to separate the 4 quadrants of the Empathy Map (says, does, thinks and feels), which could have had a positive influence at the time of applying the Brainstorming.

6. Threats to validity

It is necessary consider certain issues that may have threatened the validity of the experiment:

- External validity may be threatened when experiments are performed with students, as doubts have been raised regarding the representativeness of the subjects with respect to software professionals. Despite this, the tasks to be performed did not require real world experience, and we are, therefore, of the opinion that this quasi-experiment could be considered appropriate, as suggested in literature [2]. There are no threats related to the material used, as to the real systems were employed.
- Threats to internal validity are mitigated by the design of the quasi-experiment. In our case, both the support material and the exercise were the same under both conditions, but an additional technique was presented to one of the groups.
- Conclusion validity concerns the data collection, the reliability of the measurement and the validity of the statistical tests. Statistical tests were used to reject the null hypotheses. We have explicitly mentioned and discussed when non-significant differences were present. It is also necessary to state that the conclusion validity could be also affected by the number of observations. Further replications with larger datasets are, therefore, required to confirm or contradict the results shown herein.
- Construct validity may be influenced by the measures used to attain a quantitative evaluation of the ideas generated, the comprehension of the techniques explained, the experimentation tasks and the post-experiment questionnaire. We measured the number of ideas in order to avoid subjectivity regarding the way in which they were written, but not having proposed a template could have negatively influenced the classification process. The post-experiment questionnaire was designed using standard forms and scales.

7. Conclusions

This paper describes a quasi-experiment carried out to investigate whether the use of Empathy Maps improves the performance of the Brainstorming technique when used in the context of a requirements elicitation process. The quasi-experiment was carried out

with 72 undergraduate students enrolled on a Bachelor's degree in Computer Engineering at the Universitat Politècnica de València (UPV) in Spain. The main findings obtained are the following:

- The total number of ideas does not differ significantly between treatments, and the descriptive analysis shows that the number of ideas of requirements (technological) was lower in Group A, to which the *EM+P+B* treatment was applied. This gives rise to the theory that the Empathy Map helped the participants to be more aware of functional requirements than non-functional ones. This result was not, however, sufficiently significant to confirm the hypothesis, perhaps because the number of groups involved in the quasi-experiment were not sufficiently large.
- The descriptive analysis of the post-experiment survey shows some promising findings: despite the fact that the Perceived Utility of the Empathy Map was lower than for the other techniques, the Perceived Utility of the Brainstorming technique was higher in the *EM+P+B* treatment. This suggests that, by incorporating the evaluation of emotional and cognitive factors, the Empathy Map helped the students to carry out Brainstorming sessions that were perceived to be more useful. However, the Perceived Utility results obtained for Personas and Brainstorming had a lower standard deviation when using the Empathy Map. This could be a significant contribution for the requirements community in terms of the stability of the process that we are defining.
- It is also interesting that the perception of utility was similar when the perception of effort was greater in the *EM+P+B* treatment, which would imply that the product is valued over the effort required to carry it out.
- With regard to the sequence and connection of the techniques, although the median coincides in both treatments, the standard deviation of the results is lower for the *EM+P+B* treatment. This again is an indication of the advantage of using the Empathy Map, because it shows that there was a better relationship between the students and the technique, which contributes to the impact of its teaching in the educational field.
- We learned that not having defined a specific measure with which to evaluate the Personas technique prevented us from discovering whether the Empathy Map had any influence on it. The post-experiment survey analysis allowed us to discover that the Perceived Utility of Persona technique is lower in the Empathy Map treatment, which could be interpreted as a certain level of overlapping between the two techniques such that the students found it to be redundant and, therefore, not so useful. It might, therefore, be worth comparing the two techniques in the future.

Finally, we are aware that the results obtained from this quasi-experiment are preliminary and that a further replication is required to obtain more conclusive results. We consequently plan to carry out new studies by, for example, using either Empathy Maps or Personas before the brainstorming session in different groups. Moreover, in order to avoid disparity as regards the writing of the ideas of requirements, we plan to use User stories with a common template, which are widely employed in Agile development environments. Our long-term research goal is to continue investigating the synergy of DT methods, practices, techniques and activities with the elicitation of requirements so as to provide a more solid knowledge base with empirical data that will improve the current practices of the Information Systems requirements elicitation.

Acknowledgements

The research work presented in this paper is framed within the following projects: GEMA (JCCM, Spain, SBPLY/17/180501/000293), UX-Adapt (Generalitat Valenciana, Spain, AICO/2020/113), AETHER-UCLM (MICINN, Spain, PID2020-112540RB-C42), Adapt@Cloud (MICINN, Spain, TIN2017-84550-R) and Software Development Process Research Project (Requirement Engineering research line, UNTREF, Argentina).

References

1. Basili, V., Rombach, H.D.: Towards a comprehensive framework for reuse: A reuse-enabling software evolution environment. undefined. (1988)
2. Basili, V., Weiss, D.: A methodology for collecting valid software engineering data. IEEE Transactions on Software Engineering. SE-10 728–738 (1984)

3. Brown, T.: Design Thinking, *Harvard Business Review*, https://hbr.org/2008/06/design-thinking, Accessed: September 22, 2017, (2008)
4. Costa Valentim, N.M., Silva, W., Conte, T.: The Students' Perspectives on Applying Design Thinking for the Design of Mobile Applications. In: 2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET). pp. 77–86. IEEE, Buenos Aires (2017)
5. Daun, M., Hübscher, C., Weyer, T.: Controlled Experiments with Student Participants in Software Engineering: Preliminary Results from a Systematic Mapping Study. (2017)
6. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly. 13 (3), 319–340 (1989)
7. Ferreira, B., Silva, W., Barbosa, S.D.J., Conte, T.: Technique for representing requirements using personas: a controlled experiment. IET Software. 12 (3), 280–290 (2018)
8. Ferreira Martins, H., Carvalho de Oliveira Junior, A., Dias Canedo, E., Dias Kosloski, R.A., Ávila Paldês, R., Costa Oliveira, E.: Design Thinking: Challenges for Software Requirements Elicitation. Information. 10 (12), 371 (2019)
9. Hehn, J., Mendez, D., Uebernickel, F., Brenner, W., Broy, M.: On Integrating Design Thinking for Human-Centered Requirements Engineering. IEEE Software. 37 (2), 25–31 (2020)
10. Hehn, J., Uebernickel, F.: The Use of Design Thinking for Requirements Engineering - An Ongoing Case Study in the Field of Innovative Software-Intensive Systems. Presented at the 26th IEEE International Requirements Engineering Conference (RE'18), Banff, Alberta, Canada August 23 (2018)
11. Hehn, J., Uebernickel, F., Mendez Fernandez, D.: DT4RE: Design Thinking for Requirements Engineering: A Tutorial on Human-Centered and Structured Requirements Elicitation. In: 2018 IEEE 26th International Requirements Engineering Conference (RE). pp. 504–505. IEEE, Banff, AB (2018)
12. Hehn, J., Uebernickel, F., Stoeckli, E., Brenner, W.: Designing Human-Centric Information Systems: Towards an Understanding of Challenges in Specifying Requirements within Design Thinking Projects. 12
13. Jarke, M., Loucopoulos, P., Lyytinen, K., Mylopoulos, J., Robinson, W.: The brave new world of design requirements: Four key principles. In: International Conference on Advanced Information Systems Engineering. pp. 470–482. Springer (2010)
14. Kahan, E., Genero, M., Oliveros, A.: Challenges in Requirement Engineering: Could Design Thinking Help? In: Piattini, M., Rupino da Cunha, P., García Rodríguez de Guzmán, I., and Pérez-Castillo, R. (eds.) Quality of Information and Communications Technology. pp. 79–86. Springer International Publishing, Cham (2019)
15. Kahan, Ezequiel, Oliveros, Alejandro, Genero, Marcela: A systematic mapping study on the application of design thinking in requirement engineering. submitted to JUCS. (2021)
16. Kelley, T., Kelley, D.: Creative Confidence: Unleashing the Creative Potential Within Us All. Currency, New York (2013)
17. Köppen, E., Meinel, C.: Empathy via Design Thinking: Creation of Sense and Knowledge. In: Plattner, H., Meinel, C., and Leifer, L. (eds.) Design Thinking Research: Building Innovators. pp. 15–28. Springer International Publishing, Cham (2015)
18. Levy, M., Hadar, I.: The Importance of Empathy for Analyzing Privacy Requirements. In: 2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE). pp. 9–13. IEEE, Banff, AB (2018)
19. Loucopoulos, P., Karakostas, V.: System Requirements Engineering. McGraw-Hill, Inc., New York, NY, USA (1995)
20. Lyytinen, K., Loucopoulos, P., Mylopoulos, J., Robinson, W.N. eds: Design Requirements Engineering: A Ten-Year Perspective: Design Requirements Workshop, Cleveland, OH, USA, June 3-6, 2007, Revised and Invited Papers. Springer-Verlag, Berlin Heidelberg (2009)
21. Micheli, P., Wilner, S.J.S., Bhatti, S.H., Mura, M., Beverland, M.B.: Doing Design Thinking: Conceptual Review, Synthesis, and Research Agenda: Doing Design Thinking. J Prod Innov Manag. 36 (2), 124–148 (2019)
22. Prestes, M., Parizi, R., Marczak, S., Conte, T.: On the Use of Design Thinking: A Survey of the Brazilian Agile Software Development Community. In: Stray, V., Hoda, R., Paasivaara, M., and Kruchten, P. (eds.) Agile Processes in Software Engineering and Extreme Programming. pp. 73–86. Springer International Publishing, Cham (2020)
23. Rozante de Paula, T., Santana Amancio, T., Nonato Flores, J.A.: Design Thinking in Industry. IEEE Software. 37 (2), 49–51 (2020)
24. Ruchira Prasad, W.M.D., Perera, G.I.U.S., Jeeva Padmini, K.V., Dilum Bandara, H.M.N.: Adopting Design Thinking Practices to Satisfy Customer Expectations in Agile Practices: A Case from Sri Lankan Software Development Industry. In: 2018 Moratuwa Engineering Research Conference (MERCon). pp. 471–476. IEEE, Moratuwa (2018)
25. Souza, A., Ferreira, B., Valentim, N., Correa, L., Marczak, S., Conte, T.: Supporting the teaching of design thinking techniques for requirements elicitation through a recommendation tool. IET Software. 14 (6), 693–701 (2020)
26. Souza, A.F., Ferreira, B., Valentim, N., Conte, T.: An experience report on teaching multiple design thinking techniques to software engineering students. In: Proceedings of the XXXII Brazilian Symposium on Software Engineering - SBES '18. pp. 220–229. ACM Press, Sao Carlos, Brazil (2018)
27. Vetterli, C., Brenner, W., Uebernickel, F., Petrie, C.: From Palaces to Yurts: Why Requirements Engineering Needs Design Thinking. IEEE Internet Computing. 17 (2), 91–94 (2013)
28. Wohlin et Al: Experimentation in Software Engineering | Request PDF, *ResearchGate*, https://www.researchgate.net/publication/279355975_Experimentation_in_Software_Engineering, Accessed: September 09, 2020, (2012)