Could the utilization of Empathy Maps and Personas enhance the requirement elicitation process?

Ezequiel Kahan¹, Emilio Insfran², Marcela Genero³, Alejandro Oliveros¹

¹Universidad Nacional de 3 de Febrero, Buenos Aires, Argentina ²IUMTI - Universitat Politècnica de València, Valencia, Spain ³University of Castilla-La Mancha, Ciudad Real, Spain ekahan@untref.edu.ar,einsfran@dsic.upv.es,marcela.genero@uclm.es, aoliveros@untref.edu.ar

Abstract. There is a growing interest in the use of Design Thinking (DT) to enrich requirements elicitation processes. We carried out a quasi-experiment to compare the Effectiveness of Brainstorming sessions in terms of the number of ideas of requirements generated, and the Identified Stakeholders, complementing the Brainstorming sessions with two of the most widely used DT techniques: Empathy Maps and Personas. Therefore, we consider three treatments: Personas + Brainstorming, Empathy Maps + Brainstorming, and Brainstorming alone (control group). The quasi-experiment was carried out with 74 students enrolled in the Bachelor of Computer Engineering course at the Universitat Politècnica de València in 2021. The results reveal a statistically significant effect on Identified Stakeholders when using Empathy Maps. Descriptive analysis also shows an increase in Identified Stakeholders when using Personas, and in the Effectiveness of Brainstorming sessions when used together with Empathy Maps or Personas. There is also a variation in the type of ideas, with the number of functional ideas being higher, and that of non-functional ideas being lower when Empathy Maps are used. These results seem to indicate that Brainstorming sessions are improved when complemented with Empathy Maps or Personas techniques. Nonetheless, we still lack enough evidence to recommend either Personas or Empathy Maps, necessitating further validation.

Keywords: Requirement Engineering, Requirements elicitation process, Design Thinking, Empathy Map, Personas, Quasi-experiment.

1 Introduction

The first work that linked Design Thinking (DT) to Requirements Engineering (RE) appeared more than nine years ago [1] and, since then, the interest in this topic has been increasing. Nowadays, there are several studies demonstrating the potential of applying DT in synergy with RE, in particular with the requirements elicitation phase [2–5]. Although these studies show that it would be possible to improve the RE process by the application of DT [6], there is insufficient evidence as to which DT techniques are more

Cadernos do IME - Série Informática e-ISSN: 2317-2193 (online) DOI: 10.12957/cadinf.2024.80372 appropriate or yield better results. As Brainstorming is often used as an individual elicitation technique [7], previous research would suggest that it is possible to improve the

Effectiveness of the ideas of requirements generated during a Brainstorming session, complementing it with other user-centered and empathy-driven ideas [5]. As already stated in previous work, empathy is a key feature of the requirements elicitation process and also of DT in general [5, 8]. Empathy is a concept that includes both the *involuntary act* of feeling sympathy for someone else and the *cognitive act* of placing oneself in another's position and adopting their perspective [9]. It is the attempt to reconstruct the specific perspective of another person and how they perceive the situation. Although empathy development occurs in all stages of DT, it is particularly relevant during the "Empathize" stage, which is usually the first stage of the process.

In an earlier quasi-experiment, we proposed a requirements elicitation process that included an empathy stage [5]. The independent variable of interest was the elicitation technique used with two treatments 1) Use of Empathy Map technique, then Personas technique and then, having a Brainstorming session (Treatment EM+P+B), and 2) Use of Personas Technique and having a Brainstorming session (Treatment P+B). The dependent variable was *Effectiveness*, measured as the quantity of ideas of requirements (QIR) generated. Ideas of requirements is a concept which we developed from an initial approach by Emilio Insfrán, one of the authors of this study, used in the teaching of RE and are predecessors to the requirements. The ideas of requirements start from needs and objectives that are not necessarily adequately formulated, which can be linked to Loucopoulos's idea of starting from individual statements of informal and confusing requirements [9]. Several interesting insights emerged from this quasi-experiment, such as an increase in the number of functional ideas of requirements, and an improvement in the perception of usefulness of the Brainstorming technique when using Empathy Maps. Some of the lessons and limitations of this quasi-experiment were: 1) not having defined a specific treatment to evaluate the Empathy Maps technique alone prevented from discovering if it contributed more or less to the Brainstorming session; 2) the subject's perception revealed that the Perceived Utility of Personas was lower in the Empathy Maps and Personas treatment (first treatment) - this could be interpreted as a certain level of overlapping between Empathy Maps and Personas; and 3) after analysing the ideas of requirements generated by the participants, several inconsistencies, contradictions, or reiterations were found. This showed the need to provide the participants with a template for specifying the ideas of requirements during the generation phase of the Brainstorming session.

Considering the above limitations, and the need for more evidence about the usefulness of DT techniques for the requirements elicitation process, we proposed a new quasi-experiment, in which the following changes were introduced: 1) Empathy Maps and Personas techniques, were separated into two different treatments to compare the contribution of these techniques when used separately. In addition, a third treatment was added as a control group, in which no other techniques were applied prior to the Brainstorming session; 2) An analysis of the identified stakeholders was included as well. This information is valuable when analysing the quality of the ideas of requirements generated, because more accurate stakeholders identification allows to get a more complete understanding of the needs of the system. The quantity of stakeholders was consider as additional dependent variable; 3) given the popularity of User Stories (US) in the agile software development [10]-[11], it was decided in this new quasi-experiment to provide US as the reference template for specifying the ideas of requirements. The rest of the quasi-experiment conditions, including instructions and supporting materials were the same of our previous quasi-experiment.

This paper presents the results of this new quasi-experiment carried out to evaluate the effect in the *Effectiveness* and in the *Identified Stakeholders* of Brainstorming sessions when using them together with Empathy Maps or Personas. *Effectiveness* was measured in terms of the quantity of ideas of requirements obtained in the Brainstorming session. We also measured the quantity of stakeholders identified by the students and the distribution of ideas among them. The quasi-experiment was conducted with a group of 74 students enrolled on a Bachelor's degree course at the Universitat Politècnica de València in October 2021.

This work extends a paper presented at the Workshop on Requirement Engineering (WER 2023) [12]. The main contributions of the current paper are: 1) The related work has been updated and analyzed in more detail, 2) The section on the Analysis of Effectiveness has been revised, separating outcomes of current quasi-experiment from the comparison with the previous one and 3) A post-experiment questionnaire has been added to collect feedback regarding subjects' perceptions, considering variables proposed in the TAM (Technology Acceptance Model) [20].

The rest of this paper is structured as follows: Section 2 offers an overview of the related work. Section 3 introduces the main characteristics of the quasi-experiment, while Section 4 presents the data analysis and interpretation collected during this quasi-experiment. Additionally, Section 5 discusses the threats to validity. Finally, in Section 6, conclusions are presented along with suggestions for future work

2 Related work

The interest in employing DT techniques for requirements elicitation has grown in the Information System and Software Engineering field in recent years, as evidenced by the fact that several secondary studies have appeared on the subject [11, 13]. However, the empirical studies that evaluate the contribution of usual techniques in DT in RE, such as Empathy Maps, or Personas, are still very scarce. Searching Scopus for the following search string "("experiment*" OR "empiric*" OR "survey" OR "case study" OR "action research") AND ("Empathy Map" OR "persona") AND Requirement", only the following four papers were found:

- Canedo et al [14] explores the utilization of Journey Maps and Personas in requirements elicitation, analyzing their use, benefits, drawbacks, and challenges in both academic literature and industry. The study encompasses a systematic literature review and a survey involving 52 practitioners, including software developers, users, and managers. The research identifies 24 primary studies addressing the use of Journey Maps and Personas in software requirements elicitation. Most survey participants affirm that these techniques enhance understanding of requirements, foster integration, collaboration, and knowledge sharing within software development teams. Overall, the findings suggest that the majority of surveyed practitioners perceive Journey Maps and Personas as effective tools for comprehending software requirements in the development process.
- Teixeira et al. In [15] the use of the Lean Persona technique with 21 software professionals is investigated. They carried out a comparison to see whether the startup professionals use the technique in a different way from the established company professionals. Results revealed that the professionals used the technique for similar purposes and wrote up UX-related requirements in different levels of abstraction.
- Costa et al. [16] carried out an exploratory case study with 17 undergraduate Computer Science students with the aim of discovering: "What are the percep-

tions of students regarding learning DT?". 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 techniques very useful but stated that more training time was required to carry out the case study.

• Ferreira et al. [17] presents a controlled experiment carried out with 37 Computer Science undergraduate students in order to compare two Personas-related techniques: traditional Personas and PATHY. The authors analysed the efficiency of the techniques and the participants' perceptions of their use. PA-THY generated more relevant characteristics for the application design than did the technique that followed the traditional description. It was also more efficient as regards creating Personas.

Table 1 summarizes the main characteristics and findings of these four studies. Upon analyzing these studies, it can be concluded that the existing evidence consists of isolated empirical studies on various DT techniques, differing from the objective pursued in the current quasi-experiment (refer to Section 3.1). This experiment is part of a long-term investigation, the initial findings of which, as mentioned in Section 1, were presented in [5] and [17]

Referen-	Methodology	Goal	Participants	Main Findings
ce				
Canedo et al. [12]	Systematic lit- erature review and survey.	Explore utiliza- tion of Journey Maps and Per- sonas in re- quirements elicitation.	52 practition- ers (software developers, users, man- agers).	 Identifies 24 primary studies. Majority of surveyed practitioners perceive Journey Maps and Personas as effective tools for understanding software requirements Enhance integration, collaboration, and knowledge sharing within software develop- ment teams.
Teixeira et al. [13]	Investigation and compari- son with 21 professionals.	Examine the use of the Lean Persona tech- nique.	Startup and es- tablished com- pany profes- sionals.	 Professionals used the technique for similar purposes. Wrote up UX-related requirements at differ- ent levels of abstraction.
Costa et al. [maain characte- ritics 14]	Exploratory case study with 17 un- dergraduate students.	Discover stu- dents' percep- tions of learn- ing DT.	Computer Science stu- dents.	- Utilizes individual techniques (Personas, Empathy Maps) and team techniques (Brain- storming, co-creation workshop) Students find techniques useful but suggest more train- ing time is needed.
Ferreira et al. [15]	Controlled ex- periment with 37 undergrad- uate students.	Compare tradi- tional Personas and PATHY.	Computer Science stu- dents.	- PATHY generated more relevant character- istics for application design PATHY more efficient in creating Personas compared to the traditional description.

Table 1. Summary of the related work

3 Quasi-experiment description

The main characteristics of the quasi-experiment are described in the following subsections. This quasi-experiment was designed and reported by following the recommendations provided in [18]. Due to space constraints, the experimental material, guidelines to perform experimental tasks, and examples of the results of the experimental tasks performed by the subjects have been published online as an appendix [19].

3.1 Goal, variables and hypotheses

Following the GQM template [20], the goal of this quasi-experiment was to analyse *Elicitation Techniques* for the purpose of comparing them with respect to their *Effectiveness* and *Identified Stakeholders* from the point of view of requirements analysts

in the context of students enrolled on a Bachelor's degree course in Computer Engineering.

The independent variable is the *Elicitation Technique* used, taking into consideration three treatments: Personas + Brainstorming, Empathy Map + Brainstorming, and Brainstorming alone (P+B, EM+B and B, respectively).

The dependent variables were *Effectiveness* and *Identified Stakeholders*. The authors of this paper consider that a greater number of ideas of requirements generated during a Brainstorming session implies a greater Effectiveness for it. Also, the identification of a greater number of stakeholders could implies a greater degree of completeness in the requirements elicitation process [21], since they represent the holders of the needs and goals of the problem to be solved. Therefore, the following hypotheses was formulated:

- H1-0: There is no significant difference between the subjects' *Effectiveness* when using P+B or EM+B or B / H1-a: ≠ H1-0.
- H2-0: There is no significant difference between the subjects' *Identified stakehold-ers* when using P+B or EM+B or B / H2-a: ≠ H2-0.

Effectiveness and *Identified stakeholders* were measured as being the quantity of ideas of requirements generated by the students (QIR), and the quantity of different stakeholders identified by the students (QS), respectively. To define the measure QIR, since the ideas of requirements generated by the students were very different, the ideas of requirements were clustered into two categories, as is usual in RE processes:

- Functional ideas / Business-oriented (QIR-F). This category included all the ideas of requirements that describe or propose functionalities for the software system for an Animal Adoption Centre (the problem domain chosen for this quasi-experiment, which is introduced in Section 3.4).
- Non-functional ideas (QIR-NF). This category included all the ideas of requirements that describe or propose restrictions or constraints for the software system to be developed [21]. This category was sub-divided into two sub-categories: Technology-oriented ideas (QIR-NF-T), which refer to ideas of requirements that establish technological needs, and People-centered needs (QIR-NF-P), which refer to ideas of requirements where people that will use the application are the central target.

When classifying the ideas of requirements, it was necessary to define another category, called *Others*, to deal with those proposed ideas of requirements not directly related to the software system to be developed, e.g., "*Creation of tutorials on how to properly take care for animals*". Therefore, these ideas were not considered to measure the *Effectiveness*, which was calculated using the following formula: QIR = QIR-F + QIR-NF. Regarding the *Identified stakeholders*, a baseline with the stakeholders of interest was defined, according to the problem description. This baseline was used to compare with the stakeholders identified by the students (QS) and determine how complete and correct were the points of views considered by the students when proposing ideas of requirements. Although the *quality* of the ideas of requirements was not directly related to the problem domain addressed, together with the correctly identified stakeholders, may be considered an as indicator of the quality of the ideas of requirements identified in terms of coverage.

3.2 Selection of the subjects

We took a convenience sample of undergraduate students enrolled on a Bachelor's degree course in Computer Engineering the Universitat Politècnica de València. The students attended a theorical-practical course on RE during the academic year 2021-2022. The practical part of the course was divided into 3 class time shifts. This course included an introduction to and examples of use of the techniques employed in the quasiexperiment, i.e., Personas, Empathy Maps, and Brainstorming. The students 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 of a Brainstorming session when using Empathy Maps or Personas, and that Brainstorming is a groupbased technique, we set up several working groups with which to run the quasi-experiment. For this reason, each time shift was divided in groups, between 4 to 7 students, were randomly assigned by the course instructor. Three types of groups were defined: Groups A (P+B), which used Personas together with Brainstorming, Groups B (EM+B), which used Empathy Map together with Brainstorming, and finally, Groups C (B), which used only Brainstorming. It was decided that the time shift with the fewest students would be assigned to Group C (control group). Table 2 summarizes number of students and sub-groups for each group.

Groups	Techquines	# of Stu- dents	# of Sub- Groups
A (P+B)	Personas + Brainstorming	27	5
B (EM+B)	Empathy Map + Brainstorm- ing	33	6
C (B)	Brainstorming	14	2

Table 2. Group characterization

3.3 Experimental object, tasks, and design

The experimental object of the quasi-experiment describes the characteristics and principal needs of an Animal Adoption Centre, called "MODEPRAN". This description provided the subjects with an overview of and context in which to begin identifying the main stakeholders, and the scope in which to propose the ideas of requirements for the software system during the Brainstorming sessions. This domain was chosen because the participants may be familiar with the problem to address, and also because it does not have a strong technical component. The authors of this paper consider that a very unknown or highly-technical problem could influence negatively the objective of the experiment. In addition, this case has a moderate length, that can be addressed in one lab session without the need for an intensive training or explanation of the concepts to be managed. Both, the case and the support material, are the same used in [5], so it has been tested and validated in terms of clarity of the requested instructions. A betweensubject design was used, meaning that the subjects (i.e., working groups) in the quasiexperiment were assigned to different treatments, with each working group experiencing only one of the treatments. The experiment first task was divided into the three phases of a Brainstorming session: the preparation phase, in which the working groups were established, the procedure was explained, and the problem statement was reviewed; the generation phase, in which the generation of ideas of requirements was carried out; and the consolidation phase, in which the ideas of requirements were consolidated, reformulated if needed, and prioritized. This final phase was carried out partially during the class, and the rest remained as homework to be completed and delivered later as a final document summarizing the work done. The experimental task included the generation of ideas of requirements by means of a Brainstorming session using only

Personas, in the case of the Groups A (P+B), using Empathy Maps, in the case of the Groups B (EM+B), or running the Brainstorming session without any previous technique Groups C (B). Both Personas and Empathy Maps were created by the sub-groups of students themselves using the provided material. The second task was the completion of a post-experiment questionnaire to collect feedback regarding subjects' perceptions. Hereafter the three phases will be described:

- Preparation phase
 - Review the concepts on Brainstorming, User Stories, Personas and Empathy Maps using the slides of the course mentioned in Section A1 of the appendix.
 - Review the problem statement of the work (the Animal Adoption Center MODEPRAN) which is presented in Section A1.1. of the appendix.
 - Access the current MODEPRAN web site to know the actual information about the center and other web sites of animal adoption centers to collect ideas and learn about the domain.
 - Set the different working groups and choose a moderator and a secretary for the group.
 - Collaboratively build the Empathy Maps or Personas for each stakeholder (if needed). This only applies to Groups A and B, Group C, as control group, will only enumerate the set of relevant stakeholders. Each member of the group must be responsible for at least one Empathy Map or Personas.
 - Check the completeness of Empathy Maps or Personas reaching an agreement in the group, if needed.

• Generation phase

- Start the brainstorming of ideas of requirements.
- Write down the ideas on a post-it, usually 2 or 3 words and make the post-it visible to all participants.
- Each group member should be responsible for (at least) one stakeholder's viewpoint.
- Review the generated ideas and check if they are complete enough.
- Each group should have at least 40 ideas.
- Check if all stakeholders' viewpoints have been considered.

• Consolidation phase

- Review all the ideas of requirements generated.
 - Consolidate ideas (classify, group, merge, reformulate, etc.)
 - Agree on the priority of each idea of requirement (high, medium, low).
 - Discard those ideas that are considered not good enough (too generic, too ambitious, etc.)
 - Agree on the type of idea of requirement (i.e., functional, non-functional, people-centered).
- Create a Context Diagram.
- Specify with more details the identified ideas of requirements using the template for User Stories.
- Prepare a Document of Ideas of Requirements.

After the Brainstorming sessions, participants were asked to individually complete a questionnaire. It included both closed questions in order to analyse the perception-based variables of this study (PEOU, PU and ITU) and one open question that would enable the participants to express their opinion about the use of the techniques. Table 3 shows the closed questions included in the post-experiment questionnaire. The questionnaire contained questions related to students' perceptions when performing the experimental tasks using each treatment and intention to use the techniques. These questions were based on Technology Acceptance Model (TAM) [22], and measured the Perceived Ease Of Use (PEOU), Perceived Utility (PU) and Intend To Use (ITU) using a five-point Likert scale. The number of questions related to PEOU, PU and ITU were similar for Groups A (P+B) and B (EM+B), but different for each Groups C (B) because the first ones have additional techniques (Personas and Empathy Maps), and specific questions related to this technique were, therefore, added to the questionnaire.

ID	Classific- ation	Groups A	Groups B	Groups C
Q1	PEOU1	Personas was easy to learn to do	Empathy Map was easy to learn to do	-
	PEOU2	The Brainstorming was easy to learn to do	The brainstorming was easy to learn to do	The Brainstorming was easy to learn to do
Q3	PEOU3	Personas was easy to do	Empathy map was easy to do	
Q4	PEOU4	Brainstorming was easy to do	The brainstorming was easy to do	The brainstorming was easy to do
Q5	PEOU5	Personas was easy to use	Empathy map was easy to use	Persona was easy to use
Q6	PEOU6	Brainstorming was easy to use	Brainstorming was easy to use	Brainstorming was easy to use
Q7	PEOU7	Classifying Personas information in the different quadrants was dif- ficult for me	Classifying the Empathy Map information in the 4 quadrants (says, thinks, feels and does) was difficult for me	
Q8	PEOU9	It was easy to generate user sto- ries	It was easy to generate user stories	It was easy to gener- ate user stories
Q9	PU1	I think that Personas is useful to understand the stakeholder more	I think that Empathy map is useful to under- stand the stakeholder more	
Q1 0	PU2	I think using Personas reduces time and effort to obtain user sto- ries	I think using Empathy map reduces time and effort to obtain user stories	
Q1 1	PU3	I think that in general Persona is useful	I think that in general Empathy map is use- ful	I think in general Brainstorming was useful
Q1 2	PU4	I think using Personas allows more user stories to be obtained	I think using Empathy map allows more user stories to be obtained	
Q1 3	PU5	Having made Personas was use- ful for brainstorming	Having made Empathy map was useful for Brainstorming	
Q1 4	ITU1	I would recommend using Per- sonas to obtain user stories with brainstorming	I would recommend using Empathy map to obtain user stories with brainstorming	
Q1 5	ITU2	It would be easy for my being ex- pert in the use of Personas to ob- tain user stories	It would be easy for my being expert in the use of Empathy map to obtain user stories	

Table 3. Post-experiment questionnaire

The following multiple documents were defined as instrumentation, and some of them are presented in Appendix [19] due to space constraints:

• Slides of the course Requirements Engineering in which the students are enrolled. These slides contain a detailed description with examples of Brainstorming, User Stories, Personas and Empathy Maps.

- The problem statement, which includes a description of the overall characteristics of the Animal Adoption Center (MODEPRAN) and the URL of the actual web site.
- The structure of the Document of Ideas of Requirements.
- The User Story template.
- A questionnaire that the students will access after the consolidation phase of the Brainstorming session to obtain feedback about the work carried out.

3.4 Execution

The students were not aware that they were participating in a quasi-experiment. For them, this activity 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 the experiment were performed in two sessions over two weeks. The first week was the training session, whose purpose was to introduce the concepts, examples, and short exercises concerning the techniques that would then be applied in the quasiexperiment: Personas, Empathy Map, and Brainstorming. The quasi-experiment took place in the second week. During the execution, students were assigned to one of three groups A, B or C, and organised in smaller sub-groups composed of between four to seven students. The quasi-experiment was controlled, meaning that no interactions took place between the working groups. The training and experimental sessions lasted approximately three hours each. Once the quasi-experiment had finished, two of the authors of this paper classified the ideas of requirements obtained by each of the working groups in accordance with the classification introduced in Section 3.1 (i.e., functional, non-functional, others). The authors of this paper then analysed and classified each idea of requirement into one or more of the categories defined in Section 2.2, reaching a consensus when necessary. Examples of the results of the experimental tasks performed by the groups can be found in [19].

4 Data Analysis and Interpretation

In this section the data analysis and interpretation of the results obtained in the quasi-experiment is presented.

4.1 Analysis of Effectiveness

Table 4 classifies the descriptive statistics of the ideas of requirements. Groups A (P+B) and B (EM+B) shows bigger mean values than Groups C (B), which might indicate a contribution of Empathy Maps and Personas on the generation of ideas of requirements.

	Groups A (P+B)				Group	Groups B <i>(EM+B)</i>			Groups C <i>(B)</i>			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
QIR	37.4	5.68	30	46	40.67	5.20	34	47	35.5	0.71	35	36
QIR- F	31.20	8.17	22	44	36.17	3.87	32	43	30.5	2.12	29	32
QIR-NF	6.2	5.17	2	15	4.5	3.94	1	12	5	2.83	3	7
QIR-NF-T	5.2	3.70	2	11	2.67	2.87	0	8	3	1.41	2	4
QIR-NF-P	1	1.73	0	4	1.83	2.14	0	5	2	1.41	1	3

Table 4. Descriptive statistics of QIR

QIR-F is 15.9% higher and QIR-NF is 37.77% lower in Groups B than A, showing that there is also a difference in the number of functional and non-functional ideas generated between groups. For both groups, A (P + B) and B (EM + B), the number of ideas and the number of stakeholders is higher than in the control treatment. To test the hypothesis formulated, we analysed the effect of every treatment (Personas, Empathy

Maps, and Brainstorming) on the measures considered (QIR, QIR-F, QIR-NF, QIR-NF-T; QIR-NF-P) using the non-parametric Kruskal Wallis test. All these values were calculated using a standard configuration of SPSS. Is was also carried out the non-parametric Mann-Whitney U test, taking Groups A and C, and Groups B and C separately. Results obtained do not allow to reject H1-0, i.e., the techniques had no effect on the QIR. Similar results were obtained after repeating the test for each individual variable (QIR-F, OIR-NF, OIR-NF-T and OIR-NF-U), i.e., it was not possible to reject H1-0 in any of the cases. In the case of the comparison between Groups B (EM+B) and C (B), the values for QIR-F were close to the rejection condition (p-value = 0.062), with a moderate Observed Power (OP) = 0.362, which indicates a slight correlation. As in previous quasi-experiment [5], the total number of ideas does not differ significantly between the three techniques. The number of functional ideas was higher, and that of non-functional ideas was lower in Groups B (EM+B), to which Empathy Maps technique were applied. This supports the idea that Empathy Maps enabled subjects to become more aware (or emphatic) of functional requirements than non-functional ones. However, this result was not significant enough to confirm the hypothesis - this may be because the number of groups involved in the quasi-experiment was not large enough.

 Table 5. Number of ideas of requirements (QIR) by category – including data of the previous and current quasi-experiment

Even onim on t	Channe		# of Sub-	QIR		Maan	
Experiment	Groups	QIK-F	Groups	QIR-NF-T	QIR-NF-P	QIR	wiean
D	P+B	147	7	42	31	220	31.43
Previous experiment (2019) [5]	E+P+B	157	7	28	31	216	30.86
	Groups A (P+B)	156	5	26	5	187	37.4
Current experiment (2021)	Groups B (EM+B)	217	6	16	11	244	40.67
	Groups C (B)	61	2	6	4	71	35.5

Table 5 compares number of ideas of requirements by category (QIR, QIR-F, QIR-NF, QIR-NF-T and QIR-NF-P) as between the previous quasi-experiment and this current one. Although there were differences in some of the treatments between both experiments, and in the use of user stories as a format for writing ideas of requirements on current experiment, instructions and support material were similar in both. The total number of ideas generated in this quasi-experiment was, on average 21.5%, higher than in the previous one [5]. These differences can probably be explained by the use of the template of US to support the specification of the generated ideas of requirements.



Fig.1. Box-plots for variables QIRF and QIR-NF-T from the current and previous experiment

Figure 1 compares box plots for variables QIRF (functional ideas) and QIR-NF-T (technological non-functional ideas) respectively. Similar to previous quasi-experiment [5], there is an inverse relation between QIR-F and QIR-NF-T in treatments that includes Empathy Map technique, with fewer technological ideas of requirements and more functional ones. Also, there is an increase on the number of ideas of functional requirements on current experiment. A possible explanation is that the use of US helped the participants to better focus and conceptualise an idea of requirement that is relevant to the problem.

A significant aspect seems to be the fact that the number of ideas that arise in the current quasi-experiment applying the Personas technique (Groups A) or the Empathy Map (Groups B) is greater than the number of ideas generated by applying both techniques together (Empathy Maps and Personas) to run the Brainstorming session, as in the previous quasi-experiment. It can be observed that there is no positive effect in terms of the number of ideas when using both techniques (Personas and Empathy Maps) together. We believe that it can be valuable to use both techniques in cases where it is necessary to define stakeholders in a more "formal" way, something that the structure of the Personas technique could do more effectively. Empathy Maps provide a more gen-

eral view of stakeholders emphasizing their feelings and thoughts, meanwhile Personas provides a more descriptive view of the stakeholders. Moreover, there is no statistically significant difference between these techniques that would allow us to suggest or recommend one over the other.

4.2 Analysis of the post-experiment questionnaire

To construct the questionnaire, we employed three distinct groups of questions sourced from the standard TAM [22]. It is important to note that the questionnaire utilized in this experiment differed from the one employed in the previous study. Consequently, a direct comparison between the two questionnaires is not possible. In the previous study, only two treatment conditions, denoted as (EM+P+B) and (P+B), were utilized. Both treatments involved the application of the Persona technique, making it challenging to discern the specific impact of each technique. In contrast, in this current experiment, we employed three distinct treatment conditions: (P+B), (EM+B), and a control group denoted as (B). Furthermore, a single technique was utilized across all treatment conditions. As a result, the questionnaire was carefully designed to facilitate the comparison of the techniques and their effects across the different treatment groups.

The results are presented in Figure 2, which illustrates box-plots generated from the data collected during the post-experiment questionnaire. These figures are analyzed separately for each question, as detailed below.



Fig.2. Box-plots of questions Q2, Q4, Q6, Q8 and Q11 of the post experiment-questionnaire

Analysing PEOU of Brainstorming across the three groups reveals the following insights:

- Q2 The Perception of Ease of Learning is similar between Groups A (P+B) and B (EM+B), while it is slightly lower for Groups C (B). This suggests that Brainstorming can generally be regarded as an easy-to-learn technique. In our prior experiment, the treatment involving the Empathy Map (Treatment E+P+B) was perceived as more challenging to learn and use. Considering the improvements in the perception of ease of learning for all techniques in this experiment, and the fact that the previous Empathy Map treatment also incorporated the Personas technique, it suggests that the difficulty may have been more related to the presence of multiple techniques rather than the Empathy Map technique itself.
- Q4 When assessing the Perception of Ease of Execution, Brainstoming is perceived as easier to execute in Groups A P+B than the other two. In the case of Groups B (EM+B), responses exhibit greater dispersion, indicating that its ease of execution may depend on various factors.
- Q6 In assessment of the Perception of Ease of Use, Groups A again stands out as the easiest to use when compared to the other two groups.
- Comparing Q4 and Q6 with the questionnaire from our prior quasi-experiment, we observe a similar pattern. The treatment involving the Empathy Map (Treatment E+P+B) had generally been evaluated as more challenging to learn and use.
- Q8 The Perception of Ease of generating User Stories is consistent across all three groups. Since the ease of generating ideas was not evaluated in the previous experiment, direct comparison is not feasible. However, the consistent perception of ease, regardless of the treatment, may be more closely tied to the structured format of US rather than the techniques itself.

Question Q11 assesses PU of each of the primary techniques in each group (Persona, Empathy Map, and Brainstorming). Groups B (EM+B) are perceived as the most useful of the three, slightly surpassing Groups A (P+B), with the same median but an extended upper quartile. Comparing this question with the results of the Effectiveness analysis presented in section 4.1, it becomes evident that a more favorable perception of utility does not necessarily correlate with a higher number of generated ideas. This observation aligns with our previous quasi-experiment when comparing treatments with a greater number of techniques (e.g., P+B Treatment versus E+P+B Treatment in the previous study and the current P+B, EM+B, and B Treatments). In such cases, subjects tend to find the latter treatments more challenging, yet they also tend to generate more ideas. A possible explanation may be that Persona and Empathy techniques, by promoting empathy, lead to a higher volume of ideas. Additionally, the novelty of these techniques for participants, coupled with a heightened perception of difficulty, may drive more focused and concentrated usage, thereby enhancing the quantity of generated ideas.



Fig. 3. Box-plots of questions Q1, Q3, Q7, Q9, Q10, Q12, Q13, Q14 and Q15 of the post experiment-questionnaire

When comparing the Perceived Ease of Use (PEOU) between Groups A (P+B) and Groups B (EM+B), the following insights emerge:

- Q1 Perception of ease of learning is better in Groups A (P+B) than in Groups B (EM+B). Even when the value obtained is positive (score of 4 over 5), Empathy Map technique is perceived as more difficult to learn than the other two treatments.
- Q3 When evaluating the Perception of Ease of Execution, Groups A (P+B) exhibits a slight advantage over Groups B (EM+B), with a similar median but a wider distribution in the upper quartile.
- Q5 Similarly, in the assessment of the Perception of Ease of Use, Groups A (P+B) is slightly favored, with a similar median to Groups B (EM+B) but a broader spread in the upper quartile.
- Q7 Perception on the difficulty in classifying information is better in Groups B (EM+B) than in Group A (P+B). However, this comparison should be approached cautiously, as the techniques serve different purposes, and the method of classifying information differs between them.

Regarding the PU comparison between Groups A (P+B) and Groups B (EM+B), the following findings are observed:

- Q9 The Perception of Usefulness for understanding stakeholders is higher in Groups A (P+B) compared to Groups B (EM+B). This aligns with expectations, given that the Persona technique offers a more structured and descriptive approach compared to the Empathy Map.
- Q10 The Perception of Usefulness for reducing time and effort to obtain user stories shows similar results for both groups, with some dispersion in the responses. In general, both techniques, Personas and Empathy Maps, are perceived as useful. This perception correlates with the greater number of ideas generated

when using both treatments, compared to the control group, as observed in the effectiveness analysis.

- Q12 The Perception of Usefulness for obtaining user stories is comparable between the two groups, with a slight advantage for Group B (EM + B) in the upper quartile but a similar median to Groups A (P+B).
- Q13 Perception of usefulness of the technique (Persons or Empathy Map) for Brainstorming shows similar results for both treatments.

When comparing the Intention of Use (ITU) between Groups A (P+B) and Groups B (EM+B), we find the following:

• Q14 – The recommendation to use either the Persona or Empathy Map technique for obtaining user stories through brainstorming shows similar results for both treatments.

The results for Q12, Q13, and Q14 are in line with the increased number of ideas generated when using either technique compared to Treatment B.

• Q15 – The Perception of Ease of Mastery in the use of the technique (Persona or Empathy Map) for obtaining user stories is notably better for Groups A (P+B). In contrast, Group B (EM+B) is evaluated as neutral, with the lowest question-naire rating for that treatment. This outcome is consistent with the evaluation in our previous experiment, where brainstorming was generally considered a somewhat more challenging technique. This difficulty may be attributed to the need to identify the thoughts and emotions of stakeholders, a task that participants do not typically engage in.

To assess differences in responses across the three treatments, we employed the non-parametric Kruskal-Wallis test. The results are presented in Table 6, showcasing the outcomes for each measure used in the Kruskal-Wallis test. 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.

Treatment	Test	Varia- ble	P- value	ОР	ES (Mean ₁ – Mean ₂ / Standard error)	R
		Q1	.015	> 0,590	0,113	YE S
		Q3	.002	> 0,901	0,180	YE S
		Q5	.154	> 0,230	0,061	NO
	Kruskal- Wallis	Q7	.666	> 0,077	0,004	NO
Treatment D+B / Treatment EM+B		Q9	.183	> 0,243	0,034	NO
Treatment I + D / Treatment EWF D		Q10	.384	> 0,178	0,015	NO
		Q12	.696	> 0,178	0,003	NO
		Q13	.734	> 0,050	0,002	NO
		Q14	.392	> 0,060	0,014	NO
		Q15	.068	> 0,209	0,064	NO

 Table 6. Kruskal-Wallis results for variables Q1 to Q15

Treatment	Test	Varia- ble	P- value	ОР	ES (Mean ₁ – Mean ₂ / Standard error)	R
	Kruskal- Wallis	Q2	.594	> 0,100	0,016	NO
Tractment D D / Tractment EM D		Q4	.338	> 0,300	0,035	NO
/ Treatment B		Q6	.306	> 0,050	0,041	NO
		Q8	.687	> 0,093	0,013	NO
		Q11	.068	> 0,051	0,067	NO

The obtained results provide evidence to reject the null hypothesis for Groups A (P+B) vs. Groups B (EM+B) concerning variable Q1, as the p-value stands at 0.015, which is below the significance threshold of 0.05. This suggests that the perception of ease of learning is significantly better when utilizing Personas technique as opposed to Empathy Maps.

Similarly, the results also support the rejection of the null hypothesis for Groups A (P+B) vs. Groups B (EM+B) with regard to variable Q3, as the p-value is 0.002, falling below the 0.05 significance level. This implies that the perception of ease of execution is significantly more favorable when employing the Persona technique compared to Empathy Maps. Importantly, these findings align with the outcomes of our prior experiment [5], thus reinforcing the statistical significance of our observations in this study.

4.3 Analysis of Identified Stakeholders

To assess the *Identified Stakeholders* of the ideas of requirement, a baseline against which to compare them was built from the explicit description of stakeholders in the material provided to the students. Authors of identified the following 6 stakeholders:

- Adopter: Someone who adopts a pet.
- Donor: Someone who gives money, food, medicines or other elements needed from the adoption centre,
- Employee: Someone who is a staff member of the adoption centre.
- Partner: Someone who makes donations regularly.
- Sponsor: Someone who makes donations for specific animals.
- Volunteer: Someone who collaborates with the activities of the adoption center without receiving financial compensation for it.

QSB variable was created to measure the number of stakeholders that the participants identified, and which coincided with the baseline established by the authors. From the analysis of the descriptive statistics shown in Table 7, it was observed that Personas and Empathy Maps techniques contribute to a greater identification of stakeholders, with Groups A identifying 40%, and Groups B 50%, more than Groups C (control group).

	Groups P+B)	Α	(Tre	atment	Groups <i>EM+B)</i>	В	(Tre	atment	Groups B)	s C	(Trea	tment
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
QS	7	1.22	5	8	7.5	1.38	5	9	5	1.41	4	6
QS B	5.4	0.89	4	6	5.5	0.84	4	6	3.5	0.71	3	4

Table 7. Descriptive statistics of QS and QSB variables

To test the hypothesis related to *Identified Stakeholders* (H2-0), the effect of each one of the treatments (P + B, EM + B, B) was analysed on the measures considered (QS, QSB) using the non-parametric Kruskal Wallis test. It was also carried out the

non-parametric Mann-Whitney U test in pairs taking Groups A (P+B) and B (EM+B), A (P+B) and C (B), and B (EM+B) and C (B) separately. The results obtained allow to reject H2-0: for Groups B (EM+B) / C (B), variable QSB, given that the p-value is 0.049, which is lower than 0.05, i.e., the Empathy Maps influenced QS. For Groups A (P+B) / C (B), even when p-value was higher than 0.05 for all the variables, in the case of QSB with a p-value = 0.068, and OP of 0.604, the result and the observed power allow us to make a slightly correlation between the treatment and the result achieved. From the above, it can be said that both Personas and the Empathy Maps helped identify the essential stakeholders, which evidences the empowerment of using these techniques in combination with Brainstorming. In the case of Groups C (B), not only was the number of stakeholders significantly lower, but also the percentage of essential stakeholders identified. Additionally, upon analysing the stakeholders found by the subjects, it seemed valuable to establish a categorisation of stakeholders in a way that could reflect some of the expected contribution of the techniques. Authors proposed grouping the stakeholders into 3 categories that reflect the relation of the interested party with the Animal Adoption Centre. The categories proposed were:

- Ideas related to Internal stakeholders (IRI): includes the ideas related to the stakeholders who are staff members of the Animal Adoption Centre. Examples of this are: Veterinarian, Employee, Manager, etc.
- Ideas related to External stakeholders (IRE): includes the ideas related to the stakeholders who are not staff member of the Animal Adoption Center. Examples of this are: Sponsor, Volunteer, etc.
- Ideas related to Internal and External stakeholders (IRIE): includes the ideas related to the stakeholders that may either be employees or people who are not part of the Animal Adoption Center. Included in this are generic stakeholders, or ideas of requirements that mention more than one stakeholder for a single idea, and where those stakeholders belong to both the categories.

It was evaluated how many of the stakeholders identified in the different groups belonged to the IRI, IRE and IRIE categories. Table 8 shows descriptive distribution of ideas for the different groups according to the proposed categories.

	Groups P+B)	А	(Tre	atment	Groups <i>EM+B)</i>	В	(Tre	atment	Groups B)	s C	(Trea	tment
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
IRI	18.20	3.11	13	21	17.33	3.50	11	21	9.50	0.71	9	10
IRE	16.00	3.81	11	21	17.83	7.30	4	24	13.50	3.54	11	16
IRI E	3.20	2.95	0	6	5.50	5.24	1	15	12.50	2.12	11	14

Table 8. Descriptive statistics of IRI, IRE and IRIE

In Groups A, the IRI category is the majority, with almost 49% of the ideas. In Groups B, although the percentage is slightly lower than the IRE category (43% vs 44%), it remains at a high value compared to Groups C. In this, the IRI category is only 27%. This seems to indicate that the techniques contribute to the identification of ideas related to internal stakeholders. In the case of Groups C, where no techniques were applied, most of the ideas are related to external stakeholders. In addition, we see that the number of ideas that are simultaneously attributable to both categories is, in the case of Groups C, 3.8 to 1 time greater than Groups A, and 2.5 to 1 times greater than Groups B. In terms of the quantity of stakeholders, this would indicate that in the case of Groups C there is a greater proportion of "generic" ideas than is attributable to either type of stakeholder, which makes us suppose that the techniques help to generate more "precise" ideas, attributable to a specific type of stakeholder.

Once again, the effect of every one of the treatments (P+B, EM+B and B) was analysed on the measures considered (IRE, IRI and IRIE), using the non-parametric Kruskal Wallis test with the three groups A, B and C, and the non-parametric Mann-Whitney U test taking pairs of groups A (P+B) and B (EM+B), A (P+B) and C (B), and B (EM+B) and C (B). Table 9 shows the results obtained for each measure employed in the Kruskal-Wallis and in the Mann-Whitney U tests.

Group	Test	Varia- ble	P- value	ОР	ES (Mean ₁ – Mean ₂ / Standard error)	R
		IRI	0.081	< 0.744	0.418	NO
A / B / C	Kruskal-Wallis	IRE	0.283	< 0.103	0.210	NO
		IRIE	0.225	< 0.335	0.298	NO
A/B	Mann-Whitney U	IRI	0.579	0.067	-0.167	NO
		IRE	0.271	0.074	-0.332	NO
		IRIE	0.437	0.05	-0.259	NO
		IRI	0.053	0.839	-0.732	NO *
A/C	Mann-Whitney	IRE	0.430	0.1	-0.298	NO
	U	IRIE	0.076	0.966	OP error) R < 0.744 0.418 NO < 0.744 0.418 NO < 0.103 0.210 NO < 0.335 0.298 NO 0.067 -0.167 NO 0.057 -0.259 NO 0.05 -0.259 NO 0.839 -0.732 NO 0.966 -0.795 NO 0.703 -0.711 YE 0.101 -0.471 NO	
D/C	Mann-Whitney	IRI	0.044	0.703	-0.711	YE S
D/U	U	IRE	0.182	0.101	-0.471	NO
		IRIE	0.182	0.318	-0.471	NO

Table 9. Kruskal-Wallis and Mann-Whitney U test results for IRI, IRE and IRIE

The results obtained allow to reject Groups B (EM+B) / C (B), variable IRI, given that the p-value is 0.044, which is lower than 0.05., i.e., the Empathy Maps had no effect on IRI. For Groups A (P+B) / C (B), even when the p-value was higher than 0.05 for all the variable, in the case of IRI with a p-value = 0.053, and OP of 0.839 and IRIE, with a p-value of = 0.076 and op of 0.966, the result and the Observed Power allows to evaluate a correlation between the treatment and the result achieved.

Summarizing, statistical analysis shows the influence of the Empathy Maps technique on the number of identified baseline stakeholders, and a slight influence of Personas. So, the contribution of the techniques to increase the identified stakeholders of the ideas generated, is positive. Likewise, comparing these values with the number of ideas generated, it could be concluded, with some caution, that the number of identified stakeholders influence positively in the number and content of the ideas generated.

In Groups A (P+B) and B (EM+B), the number of ideas around internal stakeholders was equal to or greater than the external ones, which could be explained as a contribution of the technique to empathise with these stakeholders. Therefore, the Empathy Maps and Personas techniques would seem to facilitate empathy with stakeholders in domains in which they are not experts, such as veterinarian or employee, which would then be reflected in a greater flow of ideas. This is an important result because it confirms the contribution of the techniques to help subjects to empathise. In the case of Groups C (B), there are more ideas around external stakeholders. This was an expected result, as usually people tend to find it easier to put themselves in the place of these types of stakeholders (member, adopter, or volunteer) than the internal ones (veterinarian, employee, manager, etc.). On the other hand, the number of stakeholders that simultaneously belonged to both stakeholders (IRIE variable) were, in the case of Groups B (EM+B), 3.9 times greater than Groups A (P+B), and 2.27 times greater than Groups C (B), which shows that the ideas generated by applying Brainstorming only are, in terms of stakeholders, much more generic.

5 Threats to validity

Certain issues which may have threatened the validity of the quasi-experiment must be considered [23]:

- 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 believe, therefore, that this quasi-experiment could be considered appropriate, as suggested in the literature [15]. Even that two different techniques were used and there was a third control group, the execution of a single case study could limit the scope of the conclusions. In the future, the experiment could be replicated again, incorporating additional case studies from other domains to compare if there is an effect between the techniques and the domains. The possibility of contamination between groups, whereby students in one group may have shared information with those in another, may be considered a threat to the validity of the study. However, we made special efforts to ensure that this did not occur. Even so, if it had occurred, the effect of prior knowledge on the brainstorming process is likely to be attenuated because it was conducted in groups, so the influence of any one participant was diluted.
- Threats to internal validity are to some extent mitigated by the design of the quasi-experiment. In our case, both the support materials and the exercise were the same for all the groups, but an additional technique was presented to Group A (P+B) and Group B (EM+B). Due to time constraints some parts of the experiment tasks were completed *a posteriori*, outside of the controlled environment. Although this has occurred in all groups, i.e., for all treatments, it is an aspect that may have affected the results. We will therefore take it into account in future replications and experiments.
- Conclusion validity concerns the data collection, the reliability of the measurement, and the validity of the statistical tests. Statistical tests appropriate to the type of measures of the dependent variables have been used to test the hypotheses. It has been explicitly mentioned and discussed whenever non-significant differences were found to be present. It is also necessary to state that the conclusion validity could also be 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, and the experimental tasks. The number of ideas of requirements were measure, to avoid any subjectivity as regards the way in which they were written. Since participants were asked to generate ideas of requirements that still need to be negotiated and validated with the clients, we paper understand that it is an interesting and valuable result for the requirements engineer since the expected result of the use of the Brainstorming technique as an elicitation tool is the generation of a large flow of relevant ideas of requirements but not necessarily high-quality ideas of requirements, which may be performed during the negotiation and validation of the requirements of the software system to be developed.

6 Conclusions

This paper presents the results of a quasi-experiment carried out to evaluate the improvement in the *Effectiveness*, measured as quantity of ideas of requirements generated

by the students, and *Identified Stakeholders*, measured as quantity of different stakeholders identified by the students, of Brainstorming sessions when are complemented with Empathy Maps or Personas techniques. The quasi-experiment was carried out with 74 undergraduate students enrolled on a Bachelor's degree in Computer Engineering at the Universitat Politècnica de València in October 2021.

The main findings obtained are as follows:

- Analysis of descriptive statistics reveals an increase in number of ideas of requirements generated in Brainstorming sessions when using Empathy Maps and Personas techniques even when there is no statistically significant difference. Moreover, when using Empathy Maps, there is an increase in the number of functional and a reduction in that of non-functional ideas of requirements, which evidences that the technique influences the type of idea identified, as was already stated in previous studies. This result suggest that it may be useful to complement Brainstorming sessions with the use of DT-based techniques.
- The use of a US template seems to contribute to the generation of a greater number and more precise ideas of requirements. This finding emerges from comparing the number of ideas in this quasi-experiment with the results obtained in our previous study and from the post-experiment questionnaire.
- The evaluation of the post-experiment questionnaire shows that the techniques were easy to learn. They are also perceived as useful both for the Brainstorming session and for the generation of user stories. Perception of ease of learning and perception of ease of doing are better using Personas than Empathy Maps. Both results confirm results of our previous experiment [5], achieving statistical significance on this occasion, which means that Empathy Map technique is perceived as more difficult to master than Personas technique. However, comparing this to *Effectiveness*, it would not seem to have a negative impact on generating ideas or identifying stakeholders.
- The number of stakeholders identified, according to the baseline defined by the authors of this paper, is greater when applying Personas and Empathy Maps as opposed to using Brainstorming alone, with a statistically significant difference in favour of Empathy Maps. This result must be evaluated with caution, since the treatment groups, particularly group C (B), were very small.
- When analyzing the distribution of ideas among internal, external, and internal and external stakeholders, we found a statistically significant difference when applying Empathy Maps on the ideas among internal stakeholders' variable. When applying Personas on the internal stakeholder and internal and external stakeholders' variables, even when there is no statistical significance, the pvalue and observed power allows to evaluate a strong correlation between the treatment. This would indicate that the techniques favour the identification of stakeholders in domains in which they are not experts.
- It would not seem to be worth using both techniques together, but rather to use either Empathy Maps or Personas. However, it could not be concluded whether Empathy Maps or Personas is better in terms of *Effectiveness or Identified stakeholders* of the ideas of requirements. These results may be useful to practitioners as well as to RE and Software Engineering lecturers since there are various techniques available for requirements elicitation but very little evidence about how to combine them to improve the quantity and quality of requirements obtained.

As future work we plan to replicate this quasi-experiment to corroborate the findings and to obtain more conclusive results. We will consider using experimental ob-

jects related to different domains and larger samples. Also, we will be considering the incorporation of additional dependent variables, which allow a better evaluation of the quality of the requirements ideas, such as a measure of correctness. In addition, we want to explore whether there is any significative difference between Empathy Maps and Personas and to identify under which circumstances (e.g., type of problem domain, number or type of stakeholders, team experience, team composition) these differences might appear. This will contribute to obtaining empirical evidence of the different techniques used in Design Thinking for requirement elicitation, ensuring it doesn't rely solely on mere assumptions.

Acknowledgments

The research presented in this paper is part of the following projects: ADAGIO (Consejería de Educación, Cultura y Deportes de la JCCM, SBLPY/21/180501/000061), AETHER-UCLM (MICINN, PID2020-112540RB-C42) y Proyecto de investigación en Procesos de desarrollo de Software (línea de investigación en Ingeniería de requisitos, UNTREF, Argentina).

References

- 1. Vetterli, C., Brenner, W., Uebernickel, F., Petrie, C.: From Palaces to Yurts: Why Requirements Engineering Needs Design Thinking. IEEE Internet Computing. 17, 91–94 (2013).
- 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, 371 (2019). https://doi.org/10.3390/info10120371.
- 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). https://doi.org/10.1109/RE.2018.00074.
- 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). https://doi.org/10.1007/978-3-030-29238-6_6.
- 5. Kahan, Ezequiel, Isfrán, Emilio, Genero, Marcela, Oliveros, Alejandro: Studying the Influence of Empathy Maps on Brainstorming for Requirements Elicitation: A Quasi-Experiment, (2021).
- 6. 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 (2018).
- Pacheco, C., García, I., Reyes, M.: Requirements elicitation techniques: a systematic literature review based on the maturity of the techniques. IET Software. 12, 365–378 (2018). https://doi.org/10.1049/iet-sen.2017.0144.
- 8. Gasparini, A.: Perspective and Use of Empathy in Design Thinking. ACHI 2015 : The Eighth International Conference on Advances in Computer-Human Interactions. 49–54 (2015).
- 9. Kahan, E., Genero, M., Oliveros, A.: Refinamiento de un proceso de elicitación de requisitos basado en Design Thinking (CIbSE 2023) CIbSE 2023, https://conf.re-

searchr.org/details/cibse-2023/cibse-2023-main-conference/22/Refinamiento-de-un-proceso-de-elicitaci-n-de-requisitos-basado-en-Design-Thinking, last accessed 2023/09/07.

- 10. Hoda, R., Salleh, N., Grundy, J.: The Rise and Evolution of Agile Software Development. IEEE Softw. 35, 58–63 (2018). https://doi.org/10.1109/MS.2018.290111318.
- Meireles, M., Souza, A., Conte, T., Maldonado, J.: Organizing the Design Thinking Toolbox: Supporting the Requirements Elicitation Decision Making. Presented at the ACM International Conference Proceeding Series (2021). https://doi.org/10.1145/3474624.3477064.
- 12. Kahan, E., Insfran, E., Genero, M., Oliveros, A.: Improving the Requirement Elicitation Process using Empathy Maps and Personas: A Quasi-Experiment. In: Anais do WER23 - Workshop em Engenharia de Requisitos., Porto Alegre, Brasil (2023).
- 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, 371 (2019). https://doi.org/10.3390/info10120371.
- Canedo, E.D., Calazans, A.T.S., Silva, G.R.S., Costa, P.H.T., Masson, E.T.S.: Use of Journey Maps and Personas in Software Requirements Elicitation. International Journal of Software Engineering and Knowledge Engineering. (2023). https://doi.org/10.1142/S0218194023300014.
- Teixeira, G., Zaina, L.: Using Lean Personas to the Description of UX-related Requirements: A Study with Software Startup Professionals. Presented at the International Conference on Enterprise Information Systems April 25 (2022). https://doi.org/10.5220/0011033200003179.
- 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). https://doi.org/10.1109/ICSE-SEET.2017.10.
- Ferreira, B., Silva, W., Barbosa, S.D.J., Conte, T.: Technique for representing requirements using personas: a controlled experiment. IET Software. 12, 280–290 (2018). https://doi.org/10.1049/iet-sen.2017.0313.
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: Experimentation in Software Engineering: An Introduction. Springer US (2000). https://doi.org/10.1007/978-1-4615-4625-2.
- 19. Kahan, E., Isfrán, E., Genero, M., Oliveros, A. Appendix: Experimental Material (2023)
- 20. Basili, V., Rombach, H.D.: Towards a comprehensive framework for reuse: A reuse-enabling software evolution environment. undefined. (1988).
- 21. Sommerville, I.: Software engineering. Addison-Wesley Pub. Co., Wokingham, England; Reading, Mass. (1992).
- 22. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly. 13, 319–340 (1989). https://doi.org/10.2307/249008.
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: Experimentation in Software Engineering. Springer (2012). https://doi.org/10.1007/978-3-642-29044-2.