# End-user perceptions on social sustainability in context-aware applications: Validation of an experiment design

Franci Suni Lopez
*Universidad Nacional de San Agustín de Arequipa, Peru*
fsunilo@unsa.edu.pe

Nelly Condori-Fernandez
*Universidade da Coruna, Spain*
*Vrije Universiteit Amsterdam, The Netherlands*
n.condori.fernandez@udc.es
n.condori-fernandez@vu.nl

Denisse Muñante
*Univ. Bordeaux, ESTIA, F-64210 Bidart, France*
d.munante@estia.fr

***Abstract*—Our lives are being transformed by innovative software applications with important social, environmental, and economic implications. Social sustainability must be considered by software engineers to address human physical, emotional and social needs. In order to investigate how users understand and perceive the social sustainability of context-aware software applications, that are built based on the HAPPYNESS framework, we investigate a set of social sustainability-quality requirements from an end-user perspective. The perceived importance of social acceptance and perceived success are also measured. In this paper, we present the experiment design and the main outcomes of our pilot study conducted with the MEGSUS workshop participants.**

***Index Terms*—social sustainability, security, satisfaction, user perceptions, experiment, pilot study.**



## I. Introduction

In the recent years, researchers from requirements engineering and software architecture have contributed to the basis of the notion of software sustainability and progress towards to the design of sustainability-aware software-intensive systems (*e.g.,* [1]–[6]).

Software sustainability is commonly defined regarding four dimensions [2], [6]–[8]: economic, social, environmental, and technical. According to Condori-Fernandez and Lago [6], the *economic dimension* aims to ensure that software-intensive systems can create economic value. It is taken care of in terms of budget constraints and costs as well as market requirements and long-term business objectives that get translated or broken down into requirements for the system under consideration. The *social dimension* aims to allow current and future generations to have equal and equitable access to the resources in a way that preserves their socio-cultural characteristics and achieve healthy and modern society. The *environmental dimension* seeks to avoid that software-intensive systems harm the environment they operate in. And, the *technical dimension* is concerned with supporting long-term use and appropriate evolution/adaptation of software-intensive systems in constantly changing execution environment. Based on these definitions, Condori-Fernandez and Lago [6] identified quality requirements that contribute to the corresponding sustainability dimensions of software-intensive systems.

Software-intensive systems can be found in different domains such as health-care, telecommunication, transportation, banking, etc. It is expected that in the near future software-intensive systems will behave autonomously thanks to the continuous monitoring. In this paper, given the complexity of this kind of systems, and the social implications behind emerging wearable sensing technologies, we aim to empirically investigate the social sustainability from an end-user perspective.

To do this study, we focus on services-based software applications that can be built based on HAPPYNESS [9]. HAPPYNESS is a quality assurance framework that focuses on emotions as a new asset for continuous enhancement of quality of service (QoS) and user experience (UX). As it is shown in Figure 1, the essential part of the framework is its middleware. It requires two additional platforms to interact with: the *context-gathering platform* provides the infrastructure to configure and manage the sensors, and it eventually gathers the context information in real-time from them. The sensors hooked to this platform are diverse. But we are particularly interested in what we call emotional sensors, as they are the main input to our middleware. And, the *service*

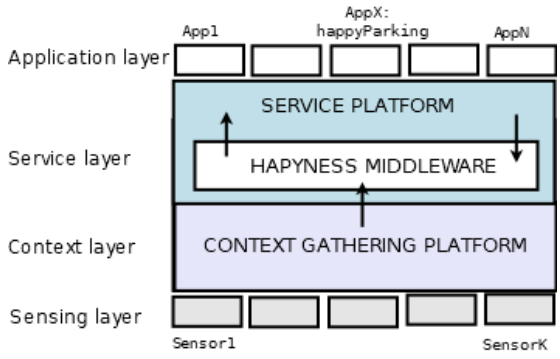


Fig. 1: An overview of the HAPPYNESS framework [9].

*platform* that provides the interface to manage and invoke available services. It can have a service composition engine and the corresponding planner, so that optimised plans can be specified according to QoS and/or UX parameters and operational service information can be obtained during their execution. Details of the HAPPYNESS *middleware* can be found in [9], [10].

The rest of the paper is organised as follows. Section II presents the scenario that illustrates the context of use of happyParking, which will be used in our experiment. Section III introduces the experiment design. The pilot design and execution is introduced in Section IV. Finally, we introduce outcome and conclusion in Sections V and VI respectively.

## II. Scenario

Frank lives in a city where the amount of parking spaces per unit is becoming scarce. Given the difficulty of finding a parking space, Frank started using a mobile application called happyParking. The application uses multiple input sources of contextual information to provide a certain degree of probability of finding a parking spot in different locations. happyParking is empowered by HAPPYNESS, and Frank uses E4Wristband[1] for monitoring emotional data at runtime. HAPPYNESS determines the actual QoS levels of happyParking services from a user perspective, increasing in this way our awareness of a potential issue with the services. It could eventually lead to actions that address the issue with the goal of improving the users experiences quality.

## III. Experiment Design

### A. Goal and Research questions

Our experiment aims to *analyse* a context-aware application, called happyParking, *for the purpose of* understanding user perceptions with respect to a set of social sustainability-quality requirements, social acceptance, and perceived success *from the viewpoint of* an end-user *in the context of* smart parking, where participants will be illustrated with two videos of two contexts of happyParking, i.e. one context related to happyParking's functionality and benefits, and the another one related to potential security vulnerabilities which can affect the use of happyParking.

From this goal, the following research question is derived: **RQ**$_1$: *How do end-users perceive the social sustainability of a context-aware software-intensive system?*

### B. Variables and Metrics

We identified two types of variables:

**Response variables:** The social sustainability is operationalised regarding: i) the perceived importance of sustainability-quality attributes, which can be measured by means of 16 items formulated in ordinal scale (see Table I). ii) Social acceptance, the term social implies individual well-being as well as the interaction between individuals [11]. With social acceptance we aim to use perception as an indicator for gathering information on what respondents think about the safety and well-being implications of different vulnerabilities of happyParking. These vulnerabilities are represented via animated-videos. Then we ask on the level of agreement on requirements that can help to mitigate the potential vulnerabilities that can affect happyParking. It can reach values from 1 (strongly disagree) to 5 (strongly agree). iii) The perceived success can measured through the Net Promoter score (NPS) [12], which can be used as an indicator of user satisfaction.

TABLE I: Qualities that contribute to Social sustainability [6]

| | **Characteristics** | **Quality attributes** |
|---|---|---|
| A1 | Security | Confidentiality |
| A2 | Security | Authenticity |
| A3 | Security | Accountability |
| A4 | Satisfaction | Trust |
| A5 | Freedom from risk | Health risk and safety risk mitigation |
| A6 | Security | Integrity |
| A7 | Effectiveness | Effectiveness |
| A8 | Satisfaction | Usefulness |
| A9 | Usability | Operability |
| A10 | Compatibility | Interoperability |
| A11 | Freedom from risk | Environmental risk mitigation |
| A12 | Usability | User error protection |
| A13 | Usability | Learnability |
| A14 | Accessibility | Accessibility |
| A15 | Usability | Appropriateness recognizability |
| A16 | Compatibility | Co-existence |

**Independent variable:** The mobile application that needs wearable sensors is identified as a variable that could affect the response variables. Our treatment is the happyParking and the E4Wristband device. Personality is another important variable identified in our study, personality will be measured using a short version of the Sixteen Personality Factor Questionnaire (16PF) [13]. For this sort version, it was selected eight factors according to technology acceptance [14].

### C. Instrumentation

The instruments to be used in the experiment are:

- A demographic questionnaire (i.e. sex, age).
- A short version of the personality test based on 16PF.
- Animated demonstration videos. The first video is used to show the main functionality of the app [2], whereas the second video is used to show specific scenarios. A specific scenario is related to one or more quality attributes. For instance a scenario-based video that shows the potential risks when security-related quality attributes were not addressed in the app [3].
- A questionnaire designed to assess the user perceptions on the social sustainability in terms of the identified response variables.

### D. Procedure

Firstly, participants are asked to complete a demographic questionnaire (1 minute) and a personality test. As the 16PF

[1]https://www.empatica.com/en-eu/research/e4/

[2]https://www.youtube.com/watch?v=iJ7Og8hTtXk

[3]https://www.youtube.com/watch?v=xDYqzzGSw8s

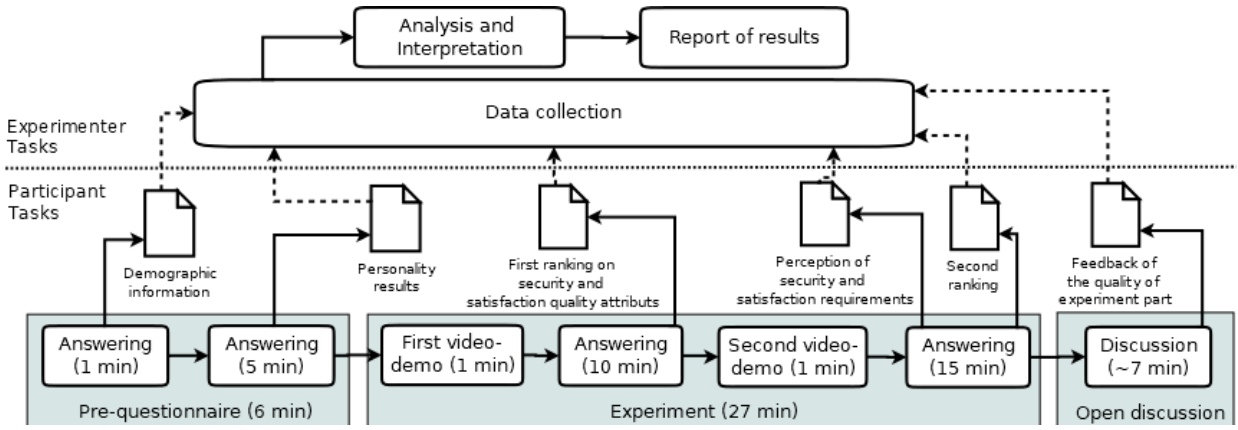


Fig. 2: An instance of the experimental procedure applied during the pilot study

test could take about 40 minutes, we plan to use a short version of it where its completion can take about 5 minutes. Then participants are asked to watch a first short video-demo (1 minute), which illustrates the context of the use of happyParking as well as its main functionality and potential benefits. Participants then are asked to complete the first part of the questionnaire to rank the social sustainability-quality attributes in order of importance (10 minutes). Next, participants watch a second animated demonstration video that illustrates a security vulnerability-related scenario (1 minute). Finally, participants are asked to complete the second part of the experiment to reveal whether the perceived importance of social sustainability-quality attributes suffer any change. Moreover, participants are asked to express their level of agreement on requirements that carry out the achievement of quality attributes. For instance, security and satisfaction-related quality attributes can be achieved by requirements that allow to mitigate potential vulnerabilities that make happyParking in danger.

As shown in Table II, the time estimated for this second part depends on the number of questions formulated for the corresponding qualities. It could oscillate between 5 and 15 minutes.

TABLE II: Estimated time for questions on risks/vulnerabilities related to social sustainability-quality requirements

| Quality characteristics | Number of questions | Estimated time |
|---|---|---|
| Security and Satisfaction | 10 questions | 15 min |
| Freedom from risk | 3 questions | 5 min |
| Compatibility | 4 questions | 10 min |
| Accessibility | 2 questions | 5 min |
| Usability | 8 questions | 10 min |
| Effectiveness | 2 questions | 5 min |

Activities of this procedure is shown in Figure 2, which also includes an additional activity, i.e. open discussion, to be carried out during the pilot study (see more details in Section IV).

### E. Threats to validity

We identified the following main threats to validity concern internal, construct and external validity defined according to [15], [16].

**Internal threats**: concern *additional factors* that may affect an observed variables. We mitigate this threat by requesting to all participants perform the experiment study in similar conditions, by using the same material (e.g. videos showing different situations when an end-user uses happyParking).

**Construct threats**: concern the *relationship between theory and observation*. In order to mitigate this threat, we do not reveal the goal of the study and the research questions before the study. Moreover, for measuring the personality of participants, we used a short version of 16PF questionnaire, which was defined and validated in the psychology field [13]. The short version consist of a subset of constructs that can have an effect on the Technology acceptance [14] (i.e. Warmth, Emotional stability, Sensitivity, Vigilance, Abstractedness, Apprehension, Openness to change, and Tension). Regarding the questions in ordinal scale (importance level) we added the option: "No opinion" to avoid forcing respondents in choosing one level of importance. Moreover, we explore the perceived importance on social sustainability-quality attributes, by formulating questions at least in two different scales.

**External threats**: concern the *generalisation* of the findings beyond the validation settings. This threat will be partially mitigated by the fact that we will invite people who are diverse in terms of personality and educational background to participate in our future full study.

## IV. Pilot design and execution

### A. Goal

We conducted a pilot of our experiment with the purpose of evaluating the feasibility, time, adverse events of our study and get feedback to improve potential deficiencies in the instruments prior to the implementation of the full study.

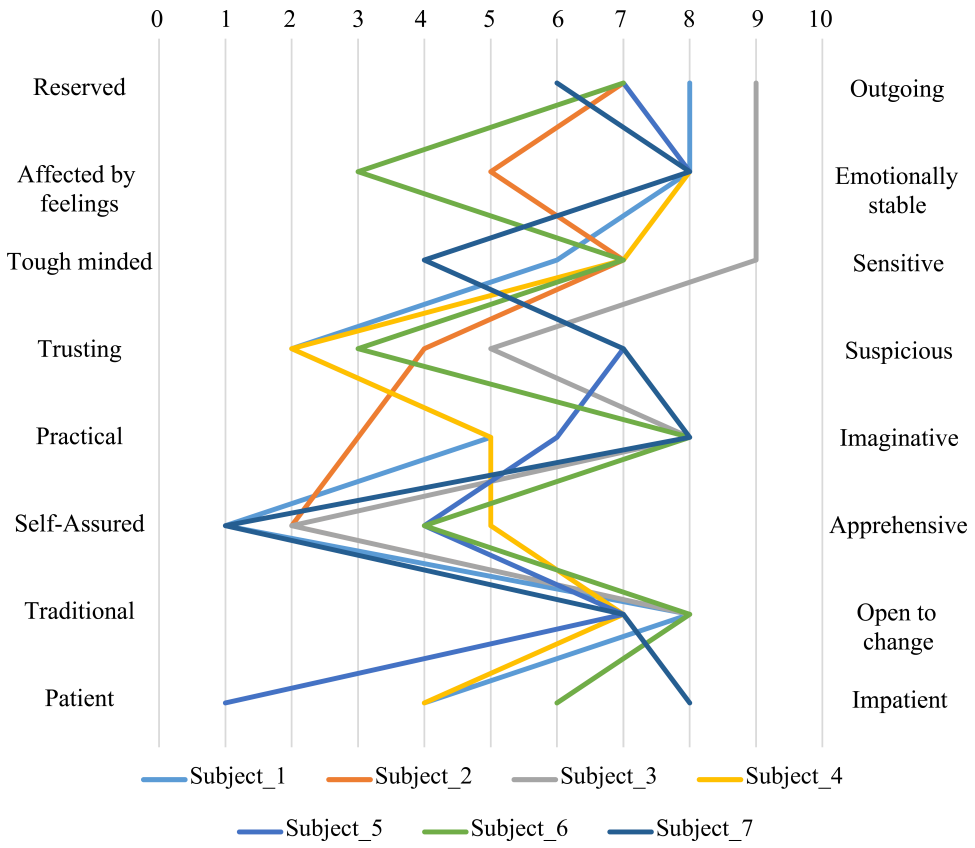


Fig. 3: Distribution of the self-report personality questionnaire of the subjects.

### B. Procedure of the pilot study

The procedure followed to conduct our pilot study is described below:

- selection of participants,
- application of the treatments (embedded videos in survey) to the pilot study,
- termination of the treatment application,
- and outcome of the pilot study.

### C. Selection of participants

To achieve the goal of our pilot study, we need potential candidates with i) high experience on running empirical studies, and ii) knowledge on topics related to software sustainability. Therefore, we consider that the attendees of the MEGSUS workshop [4] are the most appropriate candidates because of their expertise on sustainability-related research areas, as well as most of them have experience in conducting empirical studies.

Once the pilot was performed, we obtained some more details about the MEGSUS participants. In total, 7 attendees participated in the pilot. Ages of 6 participants oscillated between 29 and 59 years old, and 3 participants were females.

Additionally, according to Figure 3 that shows the responses of the self-report personality questionnaire, we clearly see that all participants can be considered mainly as *"open to change"*, and most of them tend to be *"sensitive"* (4 of 7 participants). These personality characteristics will be very helpful for getting a better interpretation of our results.

[4]http://eseiw2018.wixsite.com/megsus18

### D. Application of the treatments

Given the length of the full study, we decided to focus only on quality attributes related to security and satisfaction, *i.e.,* the six quality attributes A1, A2, A3, A4, A6 and A8 listed in Table I.

The selection of these qualities is because of the findings reported in [6], where security requirements (i.e. confidentiality, authenticity) and satisfaction (in terms of trust) were identified as high contributors.

Regarding to security, for preparing the experimental instrumentation, i.e. the second demonstration video, and questionnaires, we follow some considerations that were highlighted in [17]: users do not think they are at risk, users are not motivated in security, safety is an abstract concept, feedback and learning do not help, the evaluation of security/cost trade-off, security is a secondary task, losses perceived disproportionately to gains. Table III lists the used questions about security and satisfaction-related requirements.

The activities asked to the pilot participants and times needed for their completion are shown in Figure 2.

### E. Termination of the treatment application

Once all participants completed the questionnaire after the second video (last task), we had a more or less 7-minutes open-discussion session for collecting suggestions and comments to improve the design and instruments of the study. Part of the outcomes of this session is reported in the next section.

## V. Outcome of the pilot study

In this section, we report the results of the pilot regarding feasibility and time, identification of potential adverse events, deficiencies in the used instruments, the data analysis and visualisation plan.

### A. Adverse events and deficiencies

We identify the following adverse events that occurred during the pilot conduction:

i) Some participants had difficulties to access the online questionnaire due to the provided Google shorten URL. For sharing our survey via flyers, a possible solution to this issue is to use a QR code or create a survey link URL that is easy to type. This should not be an issue if the sharing of the questionnaire is via social media or mailing list. So, we will use multiple sharing methods for gathering a higher number of participants.

ii) As some of the participants completed the survey using their own mobile device, some items were displayed in an unreadable way. A solution that could solve this issue is to create also a mobile version of our full study.

iii) Some participants felt like the audio of the video embedded in our questionnaire was disturbing others. As this event could be a reason to reduce the non-response rate. So, we will lower the volume of the videos.

Regarding the deficiencies in our instruments, questions related to educational background were not included in the pilot because we assumed that all participants have the same educational background. However, it was warned by the participants during the open-discussion session. Thus, for a large-scale study where diverse people will be invited to participate, we plan to add questions related to the educational background as well as their experience in the use of mobile applications and wearable devices in order to have more information for getting a better understanding on end-users perceptions.

Another deficiency is related to variables used in the experiment design (see Section III). As mentioned, these variables are measured in terms of the 16 quality attributes presented in Table I. However, as new relevant quality attributes for supporting social sustainability can be identified, an extension of the sustainability model should be allowed.

In fact, during the conduction of our pilot, one of the MEGSUS participants pointed out additional issues to take in consideration for social sustainability, for instance *"users do not trust on smart apps, like happyParking, because they are built based on rules that are not transparent to them"*. From this, we can elicit a new quality requirement about *transparency*. Notice that it was not considered in TableI. Another suggestion was related to the European regulations that software designers of context-aware applications should consider. For instance, data collected from sensing devices and the way how they are stored and processed, European regulations such as the 2018 reform of EU data protection rules [5]should be taken into account. From this, another quality requirement was elicited, it is about *regulation compliance*.

Therefore, given the continuous extension of the original sustainability model, we need to provide a better management support to the experimenters. To do this, we suggest to conduct the study through the use of campaigns in a "crowd-sourcing setting" (e.g." [18]), which could facilitate i) the identification of new quality requirements using the crowd from domain experts and end-users; and ii) the selection of questions related to the social sustainability requirements that end-users would be perhaps more interested in answering.

### B. Checking the data analysis and visualisation plan

Thanks to the conduction of the pilot, we were able to check our data analysis and visualisation plan, which is very important for reporting the results of our study.

So, analysing the gathered responses from the pilot, we first verify that all subjects (7 participants) answered the questionnaire. However, once we started preparing the data, we found that one register contained incomplete data (i.e. the participant who spent more time than others, who we called $\text{Subject}_7$, answered only four items of sixteen). After checking the configuration of the questions, we realised that these items had not been configured as mandatory. As a consequence of this issue, we decided that a data register will be considered as a valid, if at least the ranking questions on the perceived importance on social sustainability quality attributes, which correspond to the second part of the pilot, were answered by the participant. In other case, the register would be discarded because it is considered as incomplete.

We also check that the data types used for the corresponding questions were correctly configured in the instrument, and the gathered data was consistent to this configuration.

Finally, we analysed and selected visualisation means for reporting the different results. It is presented as follow:

*1) The perceived importance on social sustainability-quality attributes:* To analyse this response variable, we asked 16 items about the quality attributes of social sustainability. For this analysis, we decided to use a *radar chart* that is a graphical method of displaying multivariate data. An example of this kind of chart is shown in Figure 4. The first radar chart (left side) represents the collected responses after watching the first video, whereas the second radar chart (right side) represents the answers given after watching the second video. Notice that the same questions were asked in two different moments. It was configured like this because we want to study any potential change of the importance perceptions of social sustainability-quality attributes from an end-user viewpoint. The change of perception could be provoked by the difference between the two scenarios shown in the corresponding videos.

[5]https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

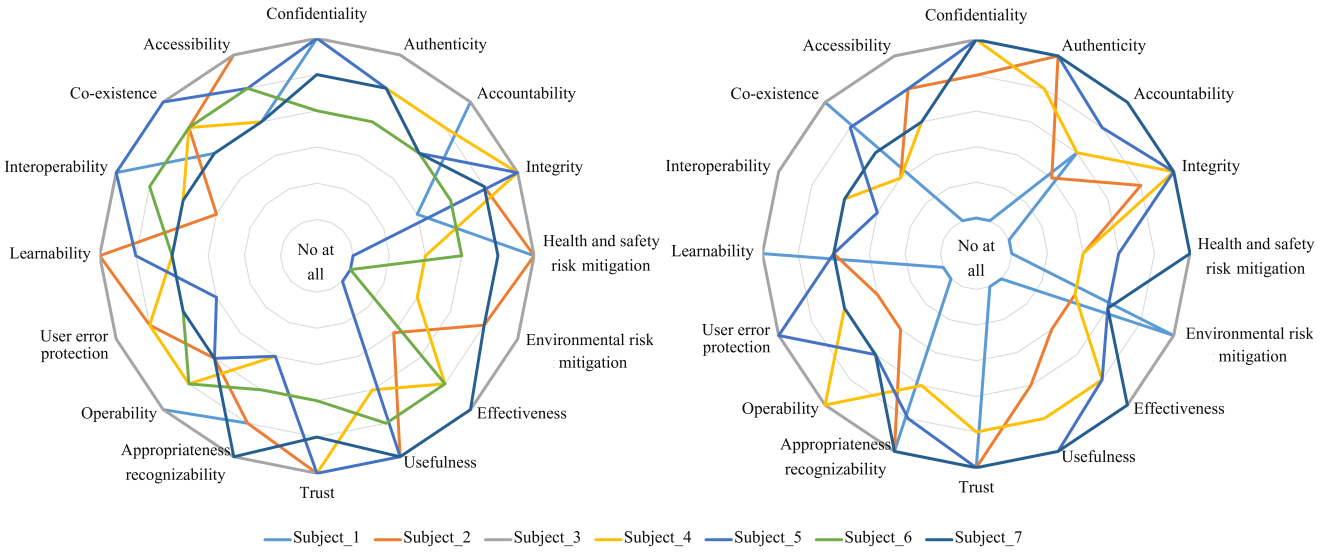


Fig. 4: Radar chart of the collected answers about the quality attributes of social sustainability. The radar on the right side represents the collected responses after watching the first video. The radar on the left side represents the answers given after watching the second video.

In Figure 4, we observe that exists a significant change of perception especially in the security-related quality attributes, where the participants rated these attributes as more important than the first time. This may be due to the impact caused by the second video that displayed the potential security vulnerabilities of happyParking. Moreover, it could also evidence that end-users do not think they are in risk unless they face to a real security attack scenario.

We plan to extend this analysis to include the demographic and personality data that is collected in the experiment to group similar participants in different clusters, for this several clustering techniques will be applied (e.g. K-means clustering technique). Moreover, each cluster could be described using Persona [19], a method to describe user profiles in order to improve the understanding about diverse user perceptions.

Once participant clusters are identified we can evaluate if there is any relation between user profiles and the tendency to perceptions changes on the social sustainability-quality attributes when different scenarios are faced.

*2) The social acceptance:* For evaluating this variable, we asked 10 questions regarding the level of agreement on requirements that can face potential vulnerabilities of happyParking. Table III presents a list of these requirements, notice that the column *quality attribute* refers to the security and satisfaction quality attribute that could be reinforced by the suggested requirement.

We use a *stacked vertical bar chart* that allows us to represent the data fragmented in percentages that correspond to different clusters in study. Figure 5 presents a summary statistics of the answers for each question and its levels of agreement. For instance, we can observe that for Q9 the 71% of the participants answered as *strongly agree* and the 29% responded as *neither agree nor disagree*. Moreover, we can note that for Q4, Q5, Q8, Q9, and Q10 the dominant answer is *strongly agree* with a percentage greater than 50%, and for questions Q1, Q6, and Q7 the principal response is *agree* that obtained greater than 40%.

TABLE III: Questions related to security and satisfaction requirements of happyParking.

| | **Question** | **Quality attribute** |
|---|---|---|
| Q1 | happyParking should register the path chose by Frank in order to trace him. | Accountability |
| Q2 | happyParking should prevent Frank in case the path proposed is dangerous, if Frank chooses this path it will be responsibility of him. | Accountability |
| Q3 | happyParking should prohibit the simultaneous access of Frank in two different devices. | Authenticity |
| Q4 | happyParking should authenticate Frank. | Authenticity |
| Q5 | happyParking should check Frank's authentication for each interaction between him and the App. | Confidentiality |
| Q6 | happyParking should ensure and protect the watch connection in order to avoid the disclosure of sensitive data. | Confidentiality |
| Q7 | happyParking should check Frank's movements and suggested paths. It should alerts Frank in case any deviation is detected. | Integrity |
| Q8 | happyParking should prevent any change on sensitive data. | Integrity |
| Q9 | happyParking should ensure Frank arrives near from his destination. | Usefulness |
| Q10 | happyParking should ensure that Frank's data is not used for other purposes beyond that finding a free parking. | Trust |

*3) The perceived success:* The perceived success was measured using the level of recommendation of happyParking. For

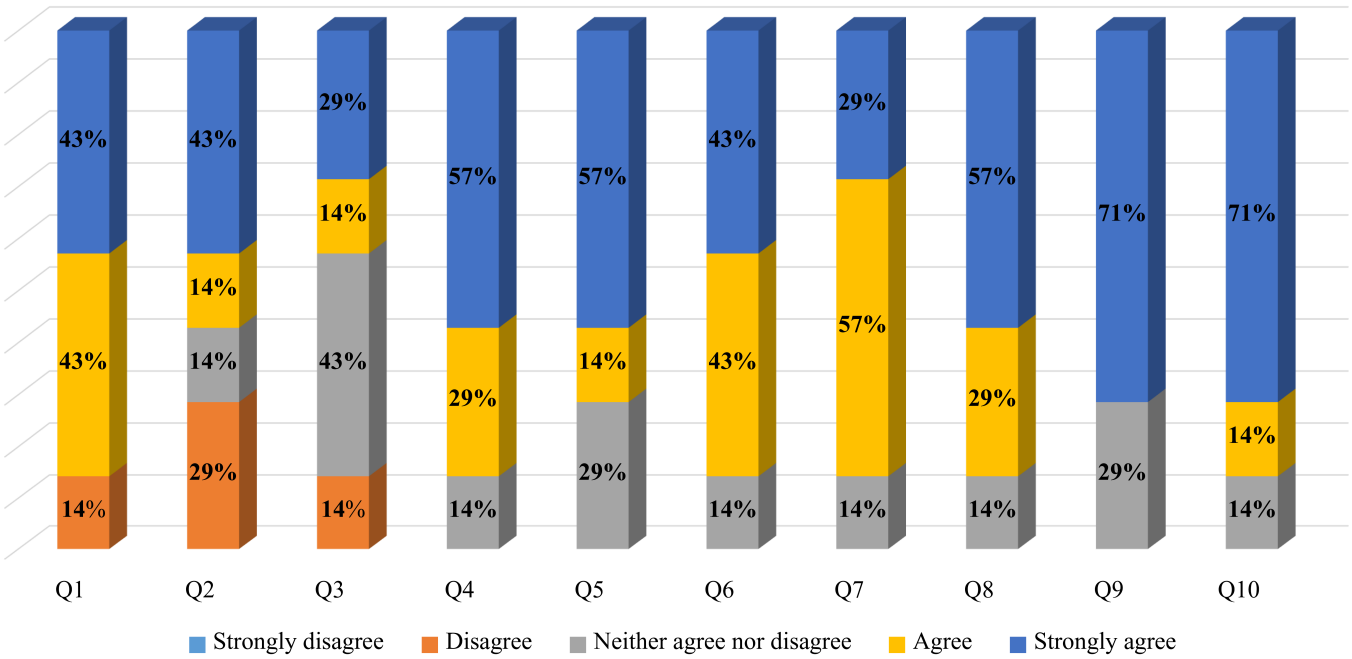


Fig. 5: Questions for evaluating the social acceptance regarding the potential vulnerabilities of happyParking.

this variable, we also asked the same question in two times. The first time corresponds after presenting the first video, and the second time corresponds after showing the second video.

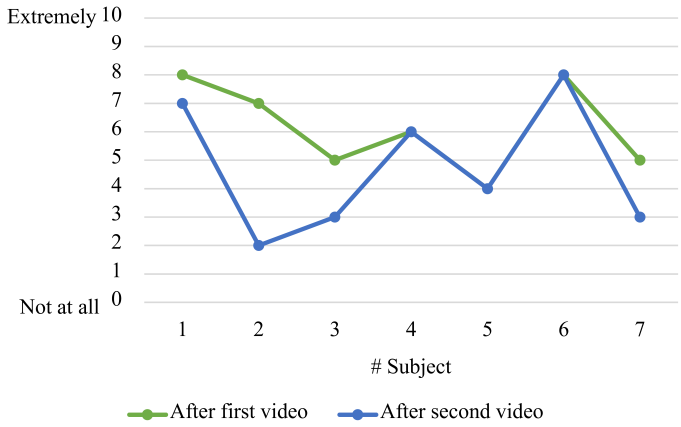


Fig. 6: Comparison between answers about the perceived success which is measured regarding the the level of recommendation of happyParking.

We calculated NPS which could be used as an indicator of user satisfaction, it can be measured in different scales, however the most used is a 0-10 scale, where 9 or 10 are promoters, 7-8 are passives, and 0-6 are detractors. The NPS score is calculated as follow:

NPS Score = % promoters - % detractors

When NPS ranges are from -100 to 0, it means that everyone is a detractor. In contrast, if NPS ranges are from 0 to +100, it means that everyone is a promoter. Consequently, after to present the first video, we get a NPS score of -57.1% (i.e. promoters = 0%, passives = 42.9%, and detractors = 57.1%). After to present the second video of the happyParking vulnerabilities, we get a NPS score of -71.4% (i.e. promoters = 0%, passives = 28.6%, and detractors = 71.4%). In other words, these results indicate the majority of participants will not recommend happyParking.

Additionally, we use a *line chart* because it allows us to visualise trends and movements over different responses. An example of this visualisation is shown in Figure 6. It compares the answers collected at the two different moments. As we can see, it shows a significant change to not at all option.

### C. Feasibility and time

Considering the actual average time needed by the pilot group (16 min in average) for completing all tasks (observe the needed time was largely covered by the planned time in Figure 2), but only focusing on two quality attributes for evaluating the social acceptance (security and satisfaction), we think that our study would be feasible if we count with the support of a crowd-sourcing platform (e.g. Amazon Mechanical Turk). This would facilitate us to segment the second part of the questionnaire per quality characteristic such is shown in Table II. Moreover, based on the experience of the last two authors in conducting empirical studies with human subjects (e.g. [18], [20]–[22] ), the ethical issues of this study will be addressed. Theses issues are regarding i) informed consent, ii) respect for anonymity and confidentiality, and iii) data privacy.

## VI. Next steps and conclusions

In this paper we have presented an experiment design for evaluating the end-user perceptions on social sustainability in context-aware applications, as well as its validation through a pilot study. In particular, we focused on the need to validate the experimental procedure and instruments, which were applied with participants of the MEGSUS workshop. The pilot has revealed some issues that have been considered for improving the study that is going to be applied in a larger-scale of participants.

As a result of the pilot execution, we have identified some potential adverse events and deficiencies in the used instruments, as well as we introduced a plan for the data analysis and its visualisation. Moreover, we have validated the feasibility of the study concluding that this study can be applied in a crowd-sourcing platform such as Amazon Mechanical Turk. So, as next steps, we will execute and validate the study in a larger and more variety audience.

## Acknowledgments

Authors would like to thank all the participants who took part of the pilot study. This work has partially received financial support from the Spanish Ministry of Economy, Industry and Competitiveness with the Project: TIN2016-78011-C4-1-R.

## References

[1] F. Albertao, J. Xiao, C. Tian, Y. Lu, K. Q. Zhang, and C. Liu, "Measuring the sustainability performance of software projects," in *IEEE 7th International Conference on e-Business Engineering, ICEBE 2010, Shanghai, China, November 10-12, 2010*, 2010, pp. 369–373.

[2] C. Becker, R. Chitchyan, L. Duboc, S. Easterbrook, B. Penzenstadler, N. Seyff, and C. C. Venters, "Sustainability design and software: The karlskrona manifesto," in *37th IEEE/ACM International Conference on Software Engineering, ICSE 2015, Florence, Italy, May 16-24, 2015, Volume 2*, 2015, pp. 467–476.

[3] S. Betz, C. Becker, R. Chitchyan, L. Duboc, S. M. Easterbrook, B. Penzenstadler, N. Seyff, and C. C. Venters, "Sustainability debt: A metaphor to support sustainability design decisions," in *Proceedings of the Fourth International Workshop on Requirements Engineering for Sustainable Systems, RE4SuSy 2015, co-located with the 23rd IEEE International Requirements Engineering Conference (RE 2015), Ottawa, Canada, August 24, 2015.*, 2015, pp. 55–53.

[4] H. Koziolek, "Sustainability evaluation of software architectures: a systematic review," in *7th International Conference on the Quality of Software Architectures, QoSA 2011 and 2nd International Symposium on Architecting Critical Systems, ISARCS 2011. Boulder, CO, USA, June 20-24, 2011, Proceedings*, 2011, pp. 3–12.

[5] C. C. Venters, C. Jay, L. Lau, M. K. Griffiths, V. Holmes, R. R. Ward, J. Austin, C. Dibsdale, and J. Xu, "Software sustainability: The modern tower of babel," in *Proceedings of the Third International Workshop on Requirements Engineering for Sustainable Systems, RE4SuSy 2014, co-located with 22nd International Conference on Requirements Engineering (RE 2014), Karlskrona, Sweden, August 26, 2014.*, 2014, pp. 7–12.

[6] N. Condori-Fernandez and P. Lago, "Characterizing the contribution of quality requirements to software sustainability," *Journal of Systems and Software*, vol. 137, pp. 289 – 305, 2018.

[7] M. Razavian, G. Procaccianti, and D. A. Tamburri, "Four-dimensional sustainable e-services," in *28th International Conference on Informatics for Environmental Protection: ICT for Energy Effieciency, EnviroInfo 2014, Oldenburg, Germany, September 10-12, 2014.*, J. M. Gómez, M. Sonnenschein, U. Vogel, A. Winter, B. Rapp, and N. Giesen, Eds. BIS-Verlag, 2014, pp. 221–228. [Online]. Available: http://www.enviroinfo2014.org/

[8] P. Lago, N. Meyer, M. Morisio, H. A. Müller, and G. Scanniello, "Leveraging "energy efficiency to software users": summary of the second GREENS workshop, at ICSE 2013," *ACM SIGSOFT Software Engineering Notes*, vol. 39, no. 1, pp. 36–38, 2014.

[9] N. Condori-Fernandez, "Happyness: An emotion-aware QoS assurance framework for enhancing user experience," in *Proceedings of the 39th International Conference on Software Engineering Companion*, ser. ICSE-C '17. Piscataway, NJ, USA: IEEE Press, 2017, pp. 235–237.

[10] N. Condori-Fernandez and F. Suni Lopez, "Using emotions to empower the self-adaptation capability of software services," in *Proceedings of the 2Nd International Workshop on Emotion Awareness in Software Engineering*, ser. SEmotion '17. Piscataway, NJ, USA: IEEE Press, 2017, pp. 15–21.

[11] G. Assefa and B. Frostell, "Social sustainability and social acceptance in technology assessment: A case study of energy technologies," *Technology in Society*, vol. 29, no. 1, pp. 63 – 78, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0160791X0600042X

[12] R. Shaw, "Net promoter," *Journal of Database Marketing & Customer Strategy Management*, vol. 15, no. 3, pp. 138–140, Jun 2008.

[13] H. Cattell and J. Schuerger, *Essentials of 16PF Assessment*, ser. Essentials of Psychological Assessment. Wiley, 2003.

[14] V. Özbek, Ümit Alnıaçık, F. Koc, M. E. Akkılıç, and E. Kaş, "The impact of personality on technology acceptance: A study on smart phone users," *Procedia - Social and Behavioral Sciences*, vol. 150, pp. 541–551, sep 2014.

[15] C. Wohlin, P. Runeson, M. Hst, M. C. Ohlsson, B. Regnell, and A. Wessln, *Experimentation in Software Engineering*. Springer Publishing Company, Incorporated, 2012.

[16] R. Wieringa, N. Condori-Fernandez, M. Daneva, B. Mutschler, and O. Pastor, "Lessons learned from evaluating a checklist for reporting experimental and observational research," in *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*, ser. ESEM '12. New York, NY, USA: ACM, 2012, pp. 157–160. [Online]. Available: http://doi.acm.org/10.1145/2372251.2372279

[17] R. West, "The psychology of security," *Commun. ACM*, vol. 51, no. 4, pp. 34–40, 2008.

[18] D. Muñante, A. Siena, F. M. Kifetew, A. Susi, M. J. C. Stade, and N. Seyff, "Gathering requirements for software configuration from the crowd," in *IEEE 25th International RE 2017 Workshops, Lisbon, Portugal*, 2017, pp. 176–181.

[19] A. Cooper, "The inmates are running the asylum," in *Software-Ergonomie '99, Design von Informationswelten, Gemeinsame Fachtagung des German Chapter of the ACM, der Gesellschaft für Informatik (GI) und der SAP AG, 8.-11. März 1999, Walldorf, Baden, Deutschland*, 1999, p. 17.

[20] F. M. Kifetew, D. Muñante, A. Perini, A. Susi, A. Siena, P. Busetta, and D. Valerio, "Gamifying collaborative prioritization: Does pointsification work?" in *25th IEEE International Requirements Engineering Conference, RE 2017, Lisbon, Portugal, September 4-8, 2017*, 2017, pp. 322–331.

[21] N. Condori-Fernández, J. I. Panach, A. I. Baars, T. Vos, and Óscar Pastor, "An empirical approach for evaluating the usability of model-driven tools," *Science of Computer Programming*, vol. 78, no. 11, pp. 2245 – 2258, 2013, special section on Mathematics of Program Construction (MPC 2010) and Special section on methodological development of interactive systems from Interaccion 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167642312001451

[22] M. A. Teruel, N. Condori-Fernandez, E. Navarro, P. González, and P. Lago, "Assessing the impact of the awareness level on a co-operative game," *Information and Software Technology*, vol. 98, pp. 89 – 116, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0950584918300314