

Cruz-Benito, J., Therón, R., García-Peñalvo, F. J., Sánchez-Prieto, J. C., Vázquez-Ingelmo, A., Martín-González, M., & Martínez, J. M. (2017). Improving success/completion ratio in large surveys: a proposal based on usability and engagement. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Technology in Education. 4th International conference, LCT 2017. Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017. Proceedings, Part II* (pp. 352-370). Switzerland: Springer International Publishing. doi:10.1007/978-3-319-58515-4_28

The final publication is available at link.springer.com

Improving success/completion ratio in large surveys: a proposal based on usability and engagement

Juan Cruz-Benito¹, Roberto Therón¹, Francisco J. García-Peñalvo¹, José Carlos Sánchez-Prieto¹, Andrea Vázquez-Ingelmo¹, Martín Martín-González², Jorge M. Martínez²

¹GRIAL Research Group. Department of Computers and Automatics. University of Salamanca

²UNESCO Chair in University Management and Policy. Technical University of Madrid

{ juan cb, theron, fgarcia, josecarlos.sp,
andrea.vazquez } @usal.es,
{ martin.martin, jorge.martinez } @upm.es

Abstract. This paper presents a research focused on improving the success/completion ratio in large surveys. In our case, the large survey is a questionnaire produced by the Spanish Observatory for University Employability and Employment (OEEU in the Spanish acronym). This questionnaire is composed by around 32 and 60 questions and between 86 and 181 variables to be measured. The research is based on the previous experience of a past questionnaire proposed by the OEEU composed also by a large amount of questions and variables to be measured (63-92 questions and 176-279 variables). After analyzing the target population of the questionnaire (with the target population of the previous questionnaire as reference) and reviewing the literature, we have designed 11 proposals for changes in the questionnaire that could improve users' completion and success ratios (changes that could improve the users' trust in the questionnaire, the questionnaire usability and user experience or the users' engagement to the questionnaire). These changes are planned to be applied in the questionnaire in two main different experiments based on A/B test methodologies that will allow researchers to measure the effect of the changes in different populations and in an incremental way. The proposed changes have been assessed by five experts through an evaluation questionnaire. In this questionnaire, researchers gathered the score of each expert regarding to the pertinence, relevance and clarity of each change proposed. Regarding the results of this evaluation questionnaire, the reviewers fully supported 8 out of the 11 changes proposals, so they could be introduced in the questionnaire with no variation. On the other hand, 3 of the proposed changes or improvements are not fully supported by the experts (they have not received a score in the top first quartile of the 1-7 Likert scale). These changes will not be discarded

immediately, because despite they have not received a Q1 score, they received a score within the second quartile, so could be reviewed to be enhanced to fit the OEEU's context.

Keywords. Human-Computer Interaction, HCI, Online Survey, Online Questionnaire, Usability, User Experience, Engagement, Trust, A/B Test.

1 Introduction

The collection of information by questionnaires and interviews is one of the most well-known and currently used methods to get users' opinions, both in the physical and digital environments.

It is common in many websites to have a form for entering information, either as a contact point, as part of the login for the system, as part of a payment process, etc. The forms are so integrated into the web user interaction, that their importance is relativized and it is assumed that the user will complete it by the mere fact that they are faced to them regularly. However, this is not so.

Indeed, the web forms pervasivity, in recent years have triggered certain trends and user behaviors towards such information entry tools. For example, it has been proven [1] the following regarding users' behavior towards forms:

- Users rely more on websites, even being more willing to perform complex actions (at all levels), such as purchases, payments, etc.
- They protect more their information, they are less willing to disclose personal information.
- They demand better products, are less tolerant to bad forms.

During the last years a lot of work has been carried out in relation to the questionnaires, establishing that users have some reluctance to complete a form from even before to begin filling it [1]. This poses certain problems regarding the achievement of information collection objectives intrinsic to any form.

Regarding the types of users who complete forms, different profiles can be set [1]:

1. Readers: Those who read the form carefully.
2. Rushers: These users rush in and begin completing fields, reading only when they think it is necessary.
3. Refusers: These users won't have anything to do with the form.

According to the literature, and intimately related to the Social Exchange Theory [2], some authors [1] distinguish three layers in the forms: relationship, conversation and appearance.

1. The relationship of a form is based on the relationship that who asks the questions has with whom responds.
2. The conversation of a form goes from the questions that are asked, to the instructions given or to the organization of the questions according to their topic.

3. The appearance of the form is the image it displays: placement of text, graphics, areas of data entry, color, etc.

Improving these factors, such as the relationship with the user, makes it easier for the user to participate and complete his task within the questionnaire.

This paper presents a research aimed at designing and validating different changes in the context of a very large questionnaire regarding users' trust, user experience, usability and engagement with the final goal of improving the users' completion/success ratios. These possible improvements are compared with another questionnaire previously developed for the same topics and context, by means of different methodologies and approaches. To present this research, the paper will have the following structure: section 2 provides the needed context of the questionnaires and population study; section 3 presents the research goals and experiments design; the fourth section comments on the proposed changes and improvements designed by the researchers; the fifth section presents the evaluation of the proposals carried out by experts. Finally, the sixth section presents the conclusions of the paper and outlines the future work to be done regarding this research.

2 Background: The Spanish Observatory for University Employability and Employment (OEEU)

During the months of June-July 2015, the Spanish Observatory for University Employability and Employment (OEEU) contacted several thousand Spanish university graduates (133588 individuals) through the universities (48, public and private institutions) where they got their degrees in the course 2009-2010 to invite them to fill out a questionnaire [3, 4].

This questionnaire had a common part with 60 questions and 167 variables measured, in addition to 3 specific itineraries depending on the users' previous responses. The first itinerary added 3 questions and 9 measured variables more. The second one, added 24 questions and 70 variables. Finally, the third itinerary added 32 more questions and 112 variables to the common part of the questionnaire.

Therefore, the questionnaire varies between 63-92 questions and 176-279 variables depending on the itinerary that the user follows. It can be stated without doubt that the questionnaire is very extensive.

The number of users who started the questionnaire was 13006 (9.74% of the total population), of which 9617 completed it (7.20% of the total population, 73.94% of the total started questionnaires).

The descriptive data regarding the age of the participants in the questionnaire were the following (the count of users is 12109 because the birthdate data was not mandatory and not all users filled it out):

count	12109.000000
mean	32.525972
std	7.018282
min	25.000000

25 %	28.000000
50 %	30.000000
75 %	34.000000
max	80.000000

As for gender, 56.05% (7290) of the users who answered the questionnaire were women and 43.94% (5716) men. In relation to nationality, 98.54% (11672) of the users were Spanish and 1.46% (173) were foreigners.

About the users who dropped out of the questionnaire, the quartiles of the dropout rate based on the questionnaire screen where they left off were:

count	3389.000000
25 %	4.000000
50 %	5.000000
75 %	7.000000

That is, 25% of the users left on screen 4 or before, another 25% left between screens 4 and 5 of the questionnaire, another 25% between screens 5 and 7 and another 25% between screen 7 and the end (depending on the itinerary).

Now, in 2017, a process of gathering information similar to the one carried out during 2015 will be conducted again. In this case, the information to be collected is about graduates of master studies that ended their studies during the 2013-2014 academic year. For this purpose, a questionnaire composed of between 32 and 60 questions and between 86 and 181 variables to be measured has been proposed (the questionnaire has again several itineraries depending on the user's answers). Without too much analysis in detail, it can be considered that despite the differences, it is a large questionnaire and shares some of the problems of the previous one in terms of difficulties or challenges that can appear during its completion by the users.

Before sending out the questionnaires to the students, the Observatory gathers some data about students from the participant universities. Currently, on February 2017, there are collected data from 28744 people coming from 32 public and private Spanish universities. About these former students, the Observatory has the following data:

Descriptive data regarding the age of the population to which the questionnaire will be addressed:

count	28744.000000
mean	35.854370
std	15.852381
min	5.000000
25 %	28.000000
50 %	31.000000
75 %	38.000000
max	117.000000

Regarding the data about the age, obviously, the aging of the population with respect to the one of the previous questionnaire is noticed. This is normal taking into account that the required age to begin a master degree is higher than that required to access to a degree (at least on a regular basis).

Regarding the gender of the population to which the questionnaire will be addressed, 55.2% (16385) are women and 44.8% (13317) are men. In relation to nationality, this is the aspect in which the current population (graduates from master degree) is more

differentiated from the study performed with degree graduates. This time, the proportion of foreign students is greater, with 88.11% (25318) of Spanish students compared to 11.88% (3414) students with foreign nationality.

In general terms, it is possible to assume that populations (putting each of them in context) are not very different. In this sense, can be highlighted the main difference is in terms of nationality. This difference could lead to consider treating differently aspects of the questionnaire to adjust to possible cultural differences. In this case, there will be no cultural distinction when designing, presenting or performing the questionnaire. This could be considered a limitation of the study.

3 Research goals and experiments proposal

3.1 Overall research goals

The main goals of the experiment that is being designed, and that will be presented below, are:

- Study how to improve the ratio participants actually starting the questionnaire (previously, close to 9.7% of the total population).
- Study how to improve the completion rate of the questionnaire (previously 73.94% completion rate).

In addition to these fundamental goals, another objective related to the second one can be proposed; namely, to grant that in case of dropout, users have completed all possible screens of the form (obtaining by this way more information even if they leave it).

3.2 Experiments proposal

For the new version of the questionnaire, it is considered that several points can be improved compared to the questionnaire implemented in 2015 and to the ways of increasing users' participation.

To implement these improvements, it is proposed to carry out two experiments in parallel:

- A study on how to improve the invitation to graduates processes.
- A study on what improvements can be implemented in questionnaires to improve participation and completion ratios.

The key aspects of each of the studies will be discussed below, indicating the main changes to be implemented, etc.

Also, before implementing these changes to the questionnaires, in addition to being supported in part by the literature, they have been subject of experts validation through a questionnaire [5].

3.2.1 *Study about how to improve the invitation processes to graduates*

In the case of the questionnaires produced by the OEEU, it is necessary to consider a fundamental factor: the privacy of the user is a primary concern over above all else (among other reasons, due to sensitive data being handled). This project is respectful and complies to Spanish Personal Data Protection Act (LOPD), having registered the OEEU's database by the Spanish authorities to safeguard the data.

Due to the privacy restrictions imposed within the project, the Observatory does not keep data that would allow to relate a person with its information. That is, there is no information related to names, ID, exact date of birth, etc. The only exception is that the Observatory offers the option to users of including their email to get information about the investigations, or the results of the draw of some devices (Android tablets) held among the graduates who complete the questionnaire.

In view of these restrictions, and because of the e-mail -if it is obtained at all occurs at the end of the whole process- the universities are the responsible for contacting their graduates offering them to participate in the process of the questionnaires. In this contact message, universities tell graduates that there is a draw among those completing this form and provide a personal link to each student to complete the task. This invitation letter designed by the OEEU Observatory could be used or not by universities, being responsible each one of them of its use and modification.

The experiment proposal in this respect is based on sending two different invitation letters. One invitation letter will be an updated version of this text used for the previous questionnaires (updated to reflect the changes related to the new edition). The second invitation letter will change both in the textual content and visual appearance, applying some changes that will be explained in following sections of this paper (basically modifying the tone and textual content of the message, plus providing a different overall design to the message [6]).

The goal of these two different invitation letters is to send one (the old version) as invitation letter for most part of the universities. The second one (the new) will be used by universities that participated in the previous edition of the questionnaires phase to test if the changes lead to variations in the entrance and participation in the questionnaire changes over the previous edition. With this proposal, it is possible to see the effect of the changes in the invitation letter (using A/B test methodologies) considering several things:

- The context of each participating university is different (population, economic factors, etc.). Therefore, specific changes are made for universities participating in both calls for data collection.
- The population of the study has changed from the first edition of the data collection to this second (age, training, etc.). For this reason, the changes between different universities will be also validated within the same edition of the questionnaires for data collection.

In the following sections, the changes to be introduced will be discussed in depth. At any rate, these proposed changes that could be introduced in the questionnaires are limited by the various constraints of the project related to privacy (it is not possible to

use external mailing platforms, etc.) and they focus fundamentally on the issues of improving trust and relationship between the user and the entity that proposes the questionnaire (the Observatory).

3.2.2 *A study about what improvements can be implemented in questionnaires to improve participation and completion ratios*

Regarding the part of the study related to the changes in the questionnaire itself, several modifications are proposed at several levels [7].

The general approach of this study is to perform an A/B test with three variants (A/B/C). The proposal is composed by a main variant (A) that follows the outline of the previous edition questionnaire (available in Spanish in http://gredos.usal.es/jspui/bitstream/10366/127374/5/Anexos_OEEU_2015.pdf), from which we have some idea of efficiency, etc., along with two other variants (B and C) that change certain issues related to the Social Exchange Theory [2].

In general, variant B of the test refers to changes related to the relationship of the participant (who answers) and who proposes the questionnaire (first layer of theory) along with changes related to the appearance (third layer of the theory) [8, 9]. More broadly, this variant B is based on trust between the parties [10, 11], further improvements and changes with respect to user experience [12], usability [9] and interface design of the questionnaire [10, 13].

On the other hand, variant C of the test includes the proposed changes in variant B plus other changes related to the relationship between the stakeholders involved in the questionnaire (first layer of the theory) and to the conversation between them (second layer). From this point of view, variant C will focus more on issues related to user engagement [6].

In any case, the three versions of the questionnaire will maintain certain rewards offered in the previous process of data collection. For example, this time there will be again a draw of electronic devices (tablets) among those who complete the questionnaire. Also, the Observatory will continue maintaining communication with those users who want to receive the latest news of the Observatory and its research.

Regarding some factors such as age, disability, or other situations and personal contexts of users, in this case they will be obviated (except the application of general accessibility standards) because the experiment is not focused on specific aspects related to possible subgroups within the population of the study [9, 14]. It is assumed that this constitutes a limitation of the study.

The effect of the changes will be measured in two ways:

- Checking the data regarding the access ratio to the questionnaires, the completion of each part of the questionnaire and the completion ratio of the questionnaire (completed screens, dropout moments, etc.).
- Evaluating the *paradata* [15]. The *paradata* from a questionnaire are the auxiliary data that describe a process, such as response times, clicks, scroll processes, etc. In this case, the *paradata* will be related to the time it takes to the user completing the task of answering each page of questions, the time to complete the full

questionnaire, the accesses to the questionnaire, etc. These *paradata* cannot be compared with similar data from the previous round of data collection about degree graduates, since nothing similar was done in that moment.

Usually, in this kind of research, users complete another questionnaire about their opinion about how they have felt about the questionnaire, how they have been able to solve the task, etc. In this case, due to the length of the questionnaire to be completed and the nature of the project, this research will not be carried out in this way. This is a limitation as to the richness of the results that can be obtained. What researchers plan to do is to invite the students, who decide to give their e-mail voluntarily at the end of the employability and employment questionnaire, to a new specialized questionnaire on these issues.

4 Proposed improvements for the questionnaire

In this section the different improvements designed for the questionnaire are described. The design process has been driven by a literature review. This literature review comprised about 650 books, papers and technical reports. The process for selecting the literature to be reviewed was:

- Making three different queries to the Web of Science and collect the results in order to iterate in reading the titles, abstracts and full content to select those papers really relevant for the topic of this research. The three queries performed were:
 - ((*"form"*) OR (*"questionnaire"*) OR (*"survey"*)) AND *"usability"* AND *"factor"* AND ((*"web"*) OR (*"online"*))
 - online forms usability
 - ((*"web"* OR *"online"*) AND (*"questionnaire?"* OR *"form?"*) AND *usability*)This process and its results are gathered in the following spreadsheet <https://docs.google.com/spreadsheets/d/1KbOCTVBqKh3Xz5nqqQY9-ywgZ2ggYNldb3OS6SasaXk/edit?usp=sharing>. In the spreadsheet the 633 unique results retrieved from the Web of Science and their status regarding to their usage in the research regarding to each review stage are presented.
- Extracting the main references from these papers and books retrieved from the Web of Science and read them. This process lead to review another 15 papers, books, standards and technical reports. Most part of them were used in some way to design the proposals that are explained below.

Once the literature was reviewed, authors designed the improvements and changes for the questionnaire. These improvements and changes are mainly supported or inspired by the literature as well as by ISO usability guidelines and HSS (U.S. Department of Health and Human Services) guidelines [16-20]. The following subsections comment each change and measure, describing for each one its purpose, its goal, the identifier associated, etc.

The ID has been set for each proposed change related to the main application area of application within the HCI discipline; despite of that, most of them apply to more

than one area, for that reason, researchers pick the main one as base for the identifier. Table 1 explains the relationship between each change/improvement (using the IDs explained in the subsections), its relationship with HCI knowledge areas or topics and with each layer of the Social Exchange Theory used as framework for the experiments design and the research in general. The main improvement areas of each change related to HCI topics are marked in red color and bigger size.

Table 1. Relationship between each change/improvement proposed, HCI application areas and layers of Social Exchange Theory

Layer of the Social Exchange Theory Improvement area regarding HCI	Relationship	Conversation		Appearance
	Trust	Engagement	Usability / User Experience (UX)	Design
TR1	X			X
TR2	X		X	X
TR3	X			
US/UX 1			X	X
US/UX 2			X	X
US/UX 3	X		X	
US/UX 4			X	X
TR4	X	X		
EN1		X		X
EN2		X		X
EN3		X		

4.1 Proposal for the invitation letter to the questionnaires

Proposed change: *TR1. Modify the text and appearance from the invitation letter to the questionnaire.*

Dear graduate (if the University wants, can personalize the email indicating the name of the graduate):

From the University (name of the University) we ask you to take 15 minutes of your time to answer a questionnaire, which for the first time applies to all Spanish universities, to know the employment status of university graduates.

The data entered will be processed anonymously by the University (name of the University) and The Spanish Observatory for University Employability and Employment (OEEU), complying with the Spanish Personal Data Protection Act (LOPD). The aggregated results will be available on the Observatory's website (www.oeeu.org).

Among the respondents to the survey will be drawn, during the month of June 2015, 10 tablets Samsung Galaxy Tab 4.

Click here to perform the survey: «URL»

Thank you for your cooperation!

A cordial greeting,
Name of University
Contact E-Mail
Contact telephone

Spanish Observatory for University Employability and Employment
udyc@oeeu.org
913364185

Fig. 1. Invitation letter proposed by the OEEU. Text translated by the authors from [5]

In the Figure 1 the basic e-mail, designed by The Spanish Observatory for University Employability and Employment to invite the graduates in the previous edition of the data gathering process, is presented. In this edition of the data gathering process related to master graduates, the basic invitation letter text will be very similar, only changing the text to reflect the master degree of the graduates and specifying that two years ago there were another similar questionnaire that collected data from degree graduates (including also the results displayed in its web <http://datos.oeeu.org>).

Among the proposed changes are the inclusion of the university logo that sends the invitation, the inclusion of the OEEU logo, a change in design to make the questionnaire according to the colors and fonts used in other OEEU's products, and changes in the text to be perceived as a more personal invitation to the graduate. These changes are intended to improve user trust in the questionnaire and the activity of the Observatory [6, 11, 13, 21].

Figure 2 shows the proposed new design (visual and textual) for the invitation letter. As explained before, the new version will be used only by few universities to allow researchers the measuring of its effect in the graduates.



Fig. 2. Invitation letter with visual and textual changes proposed for the research.
Adapted from [5]

4.2 Proposal to amend the questionnaire for variant B

Proposed change: *TR2. Adequacy of the image to the other digital products of the Observatory.*

This change is related to modifying colors, logotypes, typography, etc. to correspond the other products of the Observatory like its website <http://datos.oeeu.org>. This change is supported by the literature as a way to enhance the users' trust in the Observatory brand and products [1, 6, 8, 11, 13, 21].

Proposed change: *TR3. Inclusion of the Observatory's logo and university's logo.*

In the same way that previous proposal, the inclusion of the OEEU logo and the university logotype can reduce the distrust of the graduate to participate. In this case, the logotype of the university will help to build trust on the questionnaire website and the OEEU logotype will help him/her to associate the product with the institution that proposes it [1, 6, 21].

Proposed change: *US/UX1. Inclusion of a progress bar in the questionnaire.*

By observing a progress bar, the user can know its progress in the task of filling the questionnaire and estimate how much effort/time he/she will need to make to complete

it. This can reduce the stress related to uncertainty about a task like an unknown questionnaire [1, 6].

Proposed change: *US/UX2. Present a visual focus animation on concrete actions.*

In this case, the web will provide a visual effect of focus to the user in that he will have always in the center of the screen the task to be solved (typically answering a question or filling an empty field), making also a defocused effect on the elements that are not fundamental to solve that task. This proposal is used in commercial questionnaire systems like <http://typeform.com/>.

The reader can access to the following URLs to check how this visual effect works: <https://drive.google.com/file/d/0BwS7cZg3riXtajJtNGhkMnlzXzg/view?usp=sharing>, <https://drive.google.com/file/d/0BwS7cZg3riXtWgk1bmlvSVB5dDg/view?usp=sharing>.

Proposed change: *US/UX3. Deactivation of control elements when an action is initiated.*

A typical example of this change is to deactivate a button in a website once it is pressed until its action is finished. This usability / user experience measure could make the user to trust on the sturdiness of the system and reduce stress situations like those where a button perform the same action several times after being pressed more than once [6].

Proposed change: *US/UX4. In related elements, instead of having smaller and more specific groupings, use some larger grouping, following the Gestalt principles on grouping.*

For example, following the proposal, the header of a table would be fixed while in the content can be scrolled up and down. It seeks to ensure that the large dimensions of analysis in some points of the questionnaire are grouped in an attempt to avoid user fatigue and reducing users' cognitive load when dealing with large tables or complex visual elements [1, 6, 22].

A visual explanation of this proposal can be observed in the following URL <https://drive.google.com/file/d/0BwS7cZg3riXtdmZqQzBHZXJVcmM/view?usp=sharing>.

4.3 Proposal to amend the questionnaire for variant C

Proposed change: *TR4. Changes in the introduction text to the questionnaire.*

In this case, a change in the text will be sought in a similar way to the modification in the invitation letter to the users. The text changes to a more personal way of addressing the user and contributing important arguments to influence a better perception on what is going to be done and improving the confidence in the questionnaire and the entity that proposes it.

The text of the previous edition is presented in the Figure 3 (the variant A will only update the data about the raffle in the text, etc.).

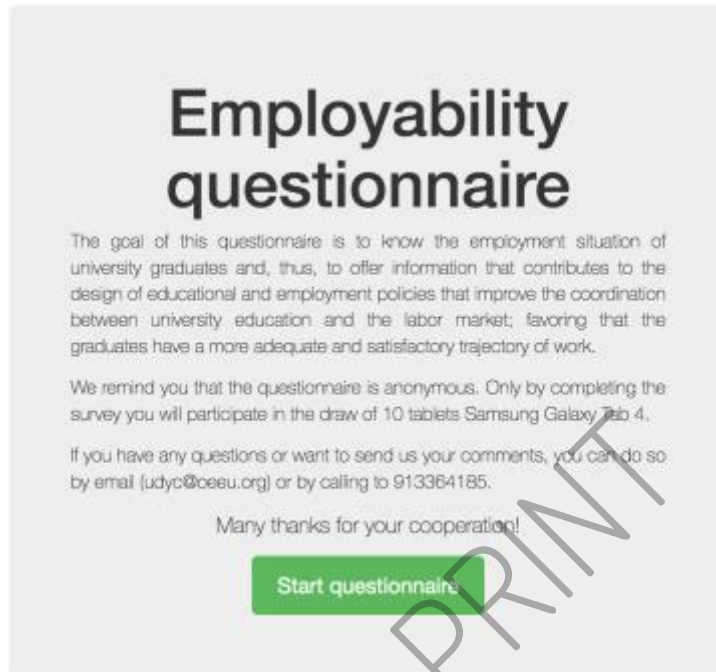


Fig. 3. Previous introduction text to the questionnaire. Translated and adapted from [4, 5]

In this case of variants B and C of the questionnaire, the introductory text would become (changing also the design and layout as commented in the proposal TR2) the displayed in the Figure 4.

Proposed change: *EN1. In the questions related to the community in which they live, change the drop-down selector for a map with the autonomous communities of Spain.*

This will allow the user to select where the user lives through clicking the corresponding one. In this case, it is sought to have visual elements different from the usual ones that allow the user to interact in different ways during the completion of the questionnaire and avoiding to suffer so much fatigue on the repetition of actions. Also, the usage of a map tries to reduce the users' cognitive load that implies the activity of reading a drop-down list of at least 20 items (autonomous communities and cities in Spain). This change is related to some authors that suggest that changing the interaction elements can affect the users' easiness to complete a task [23] and other authors that explain that the time that an user interact with elements in the form is time that the users is not thinking in dropout [6].

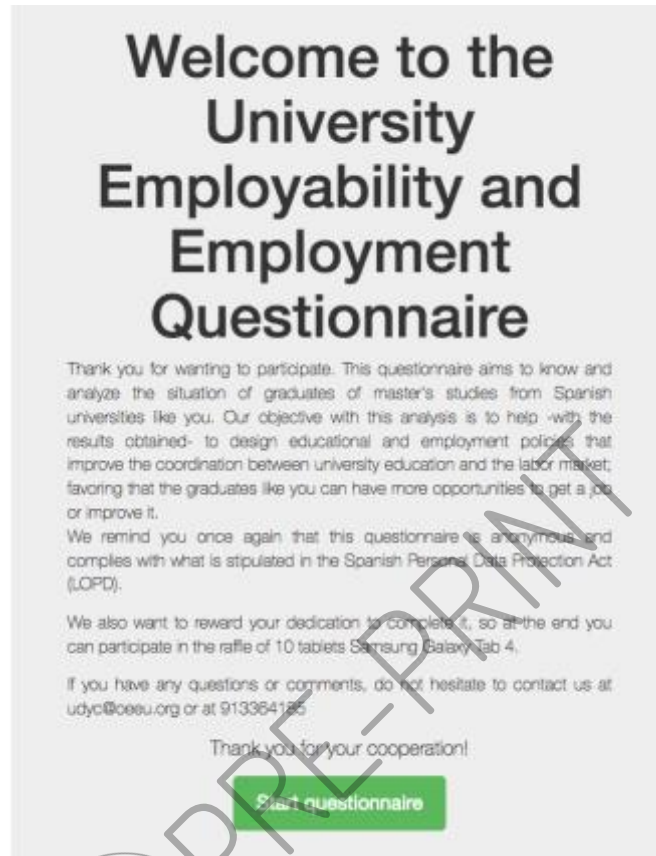


Fig. 4. Modification proposal for the introduction text to the questionnaire. Translated and adapted from [5]

Proposed change: *EN2. Inclusion of textual feedback related to user responses including information that may be relevant.*

This inclusion of textual feedback should be placed in at least three different moments of the questionnaire (i.e. after the demographic questions, after the enquiry about whether the graduated has been employed after the master degree or not, and in the final part of the chosen itinerary), regarding the different main dropout moments of the previous data collection process presented in the second section of this paper. This change requires introducing an intermediate screen between two pages of questions in the questionnaire. In this intermediate screen, information in relation to some of his/her answers enabling also comparison of their answers those provided by other users or official stats from other sources, will be provided to the user.

As an example of this kind of feedback, after the screen of the questionnaire where the user responds if he/she has ever worked and how many jobs he/she has had, in the questionnaire screen change (after pressing next), should be displayed a new screen, with only one the the following questions, should be displayed:

- If the graduate answered that he/she did not have a previously a job: “Did you know that there are XX% of graduates in your promotion who have not been able to get a job?”.
- If the graduate answered that he/she have had a job: “Do you know that the employment rate of master graduates in Spain is XX%?”.
- If the graduate answered that he/she had several jobs: “Did you know what ...? Like you there are XX% of people who have responded to this questionnaire that are in your same situation”.

Proposed change: *EN3. Inclusion of web push notifications that allow Observatory to send messages to users in order to encourage them if they leave the questionnaire before finishing.*

These notifications can only be sent if the user explicitly accepts them. The notifications will be accompanied by the link to resume the questionnaire. From a technical point of view, the notifications will be sent to Chrome, Firefox and Safari browsers on Windows, Linux and Mac OS in desktop operative systems and to Android phones with any of those browsers (estimated total market share covered by a 61-77%).

This measure can help to increase the users’ engagement as well as to try to improve the completion ratio of the questionnaires through the reinforcement.

Some examples of these kind of notifications are available (in Spanish) in [5].

Also, these web push notifications could help researchers to reach again the participants to invite them to another questionnaire to get feedback about the changes/improvements implemented finally in the form.

5 Evaluation by experts

To validate the proposals designed to improve the questionnaire and reduce the dropout ratio and increase the participation ratio, five experts were invited to evaluate the proposed measures using questionnaire. These experts were selected because all of them work usually with questionnaires from different perspectives (some of them work with questionnaires focusing on improving their usability, use them for research in several contexts, or design questionnaires as part of their day by day work).

In the following subsections, details regarding the questionnaire will be commented, as well, the results and opinions gathered from that questionnaire will be presented and discussed.

5.1 Feedback questionnaire

The assessment questionnaire completed by the experts is based on the proposal by Sánchez-Prieto et al. [24]. In it, the experts assess the relevance of each proposed change, its clarity and its importance, through a Likert scale (1-7 values). In addition the expert can comment on a qualitative way (typing comments in a textbox) any related issues to each question. Also, the questionnaire requires demographic data from the experts related to their gender, knowledge area, etc. [5] in order to characterize them.

5.2 Results and discussion

First, in the validation questionnaire, the experts completed some answers about personal information. In this case, 4 out of 5 experts (80%) were men, 1 (20%) was woman. Regarding the age, 3 out of 5 (60%) are between 41 and 50 years old, while other 2 experts (40%) are between 31 and 40 years old. Regarding their knowledge areas, 3 out of 5 (60%) are related to Engineering and Architecture, while the other 2 (40%) are related to Social and Legal Sciences. Regarding their specialization field, 3 out of 5 (60%) are related to disciplines within Computer Sciences and the other 2 (40%) are related to disciplines within Economics.

Table 2. Descriptive results from the experts' evaluation for each proposal regarding the pertinence, relevance and clarity

	Pertinence			Relevance			Clarity		
	AVG	STD	N	AVG	STD	N	AVG	STD	N
TR1	6.17	0.98	5	5.17	1.17	5	6.17	0.98	5
TR2	6.17	1.33	5	6.00	1.55	5	5.67	1.37	5
TR3	6.67	0.52	5	5.83	1.47	5	6.67	0.52	5
TR4	6.33	1.03	5	6.00	0.89	5	6.50	0.84	5
US/ UX 1	6.50	1.22	5	6.83	0.41	5	6.50	1.22	5
US/ UX 2	7.00	0.00	5	7.00	0.00	5	6.67	0.52	5
US/ UX 3	5.67	2.42	5	6.00	2.45	5	5.33	2.42	5
US/ UX 4	7.00	0.00	5	6.67	0.52	5	6.67	0.52	5
EN1	6.83	0.41	5	6.33	1.03	5	5.67	1.03	5
EN2	4.83	2.14	5	5.50	1.22	5	5.17	2.32	5
EN3	7.00	0.00	5	7.00	0.00	5	7.00	0.00	5

Table 3. Descriptive results from the experts' evaluation for each group of proposals and global assessment regarding the pertinence, relevance and clarity

	Pertinence			Relevance			Clarity		
	AVG	STD	N	AVG	STD	N	AVG	STD	N
TR	6.24	1.00	20	5.76	1.26	20	6.24	1.04	20
US/UX	6.48	1.47	20	6.62	1.32	20	6.24	1.51	20
EN	6.19	1.64	15	6.25	1.13	15	5.88	1.67	15
Global	6.27	1.30	55	6.22	1.30	55	6.11	1.42	55

Related to their responses about each proposal, as previously said, the expert had to assess the change proposal regarding the pertinence, relevance and clarity. Also in each question related to a proposal, the expert could introduce qualitative feedback through texting its opinion. Table 2 gathers the average mark, standard deviation and number of responses collected for each change/improvement proposal in terms of pertinence, relevance and clarity. Also, Table 3 gathers the same information but showing it in groupings related to the main topic associated to each change/improvement proposal (trust, usability / user experience and engagement) as well as the global average, standard deviation and number of responses collected in the assessment questionnaire.

The calculations and original responses retrieved from the experts can be checked in the sheet 2 of the following spreadsheet https://docs.google.com/spreadsheets/d/1dO72ZiHTt83UI2_cfjSd5sO1M109TXdO5rysCqgIp94/edit?usp=sharing.

In general, the average mark of the assessment in each question and grouping topic could be considered as good: most of the results are in the Q1 (score 5.5).

This Q1 score is not achieved in the proposed change EN2 (*inclusion of textual feedback related to user responses including information that may be relevant*) pertinence and clarity, TR1 (*modify the text and appearance from the invitation letter to the questionnaire*) regarding its relevance and US/UX3 proposed change (*deactivation of control elements when an action is initiated*).

Regarding the qualitative comments introduced by the experts in their feedback, the following could be highlighted:

- Comments with recommendations about visual design and layout as well as minor changes in the text of the proposed invitation letter and proposed introduction text to the questionnaire.
- Comments about the fact that many users will not know previously the OEEU's visual brand, so many graduates would not develop positive feelings regarding to trust in TR2 proposal.
- A comment regarding to US/UX3 (*deactivation of control elements when an action is initiated*) proposed change where the expert explains that he/she "is not aware about what implies this change".
- Very positive comments regarding the US/UX4 proposal (*in related elements, instead of having smaller and more specific groupings, use some larger grouping, following the Gestalt law on grouping*).
- Comments related to EN1 proposed improvement (in the questions related to the community in which they live, change the drop-down selector for a map with the autonomous communities of Spain) to include something similar for graduates that do not live now in Spain and live abroad (instead of selecting Spain autonomous communities, select countries, etc.).
- Two positive comments and another two expressing doubts about EN2 proposal (*inclusion of textual feedback related to user responses including information that may be relevant*). The positive comments explain that this change could lead to engage users by taking advantage of their curiosity. The other two explain that these extra screens and personalized feedback could break the users' trust in the data anonymization and introduce some distortions in the questionnaire.
- Some comments regarding little details that could improve the notifications. For example: the text to accept the reception of notifications should be "Yes, I accept" instead of "Ok, I accept" or introduce information about how much time will take to the user to complete the questionnaire if the he/she continues it.

In general terms, the feedback from the experts about the proposed changes/improvements for the questionnaire is very positive. Most part of the scores gathered by the experts are in the top first quartile of the scale (values 1-7), so can be

accepted “as is” to be implemented in the questionnaire of course, after a final evaluation of the convenience with the project managers and OEEU coordinators -.

On the other hand, the experts raised some doubts in other elements or certain assessment points, like the pertinence and clarity of EN2 proposal (textual feedback), the relevance of TR1 proposal (modifying the text and appearance of the invitation letter) or the relevance of the US/UX3 proposal (deactivation of control elements when an action is initiated). In these cases, all the evaluations exceeded the Q2 score (4.0 value), so still they can be considered as well perceived changes but, in any case, these should be reviewed again by the researchers, in order to improve them or discard certain proposals if there is no possible improvement for that.

Despite some of these changes that are not fully supported by experts usually are backed by other authors in the literature, researchers should follow a pragmatic approach that ensures the right application of this kind of changes/improvements for the specific case of the OEEU’s questionnaire and its context.

Also, as previously commented, all these changes and improvement proposals will be validated again with the OEEU project coordinators and OEEU project managers before implementing them in the final version of the questionnaire that will be public in April 2017.

6 Conclusions

This paper presents a research focused on improving the success/completion ratio in large surveys. In this case, the large survey is the questionnaire produced by the Spanish Observatory for University Employability and Employment and that will be publicly available for graduates of master degree in April 2017. This questionnaire is composed by about 32 and 60 questions and between 86 and 181 variables to be measured. The research is based on the previous experience of a past questionnaire proposed also by the Observatory composed also by a large amount of questions and variables to be measured.

Analyzing the target population of the questionnaire (also comparing with the target population of the previous questionnaire) and reviewing the literature, the researchers have designed 11 proposals for changes related to the questionnaire that could improve the users’ completion and success ratios (changes that could improve the users’ trust in the questionnaire, the questionnaire usability and user experience or the users’ engagement to the questionnaire). These changes are planned to be applied in the questionnaire in two main different experiments based on A/B test methodologies that will allow researchers to measure the effect of the changes in different populations and in an incremental way.

The proposed changes have been assessed by five experts through an evaluation questionnaire. In this questionnaire, researchers gathered the score of each expert regarding to the pertinence, relevance and clarity of each change proposed. Regarding the results of this evaluation questionnaire, the reviewers fully supported 8 out of the 11 changes proposals, so they could be introduced in the questionnaire with no variation. On the other hand, 3 of the proposed changes or improvements are not fully

supported by the experts (they have not received a score in the top first quartile of the 1-7 Likert scale). These changes will not be discarded immediately, because despite they have not received a Q1 score, they received a score within the second quartile. Instead of being discarded, these changes will be reviewed again by the researchers and the Observatory staff in order to adequate them to the questionnaire. If there is no possibility to adequate them to the OEEU's questionnaire context, finally they will be finally rejected.

After all this work, research and validation processes, the future work is to implement all the accepted changes and variations in the OEEU's questionnaire for graduates of master studies and study what of these changes lead to a real improvement in the completion/success ratio related to the questionnaire.

Acknowledgments

The research leading to these results has received funding from “la Caixa” Foundation. Also, the author Juan Cruz-Benito would like to thank the European Social Fund and the *Consejería de Educación* of the *Junta de Castilla y León* (Spain) for funding his predoctoral fellow contract. This work has been partially funded by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES project (Ref. TIN2016-80172-R).

References

1. Jarrett, C., Gaffney, G.: *Forms that work: Designing Web forms for usability*. Morgan Kaufmann (2009)
2. Dillman, D.A.: *Mail and internet surveys: The tailored design method*. Wiley New York (2000)
3. Michavila, F., Martín-González, M., Martínez, J.M., García-Peñalvo, F.J., Cruz-Benito, J.: *Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System*. Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15). ACM Inc., Porto, Portugal (2015)
4. Michavila, F., Martínez, J.M., Martín-González, M., García-Peñalvo, F.J., Cruz-Benito, J.: *Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)*. (2016)
5. Cruz-Benito, J., Therón, R., García-Peñalvo, F.J., Martín-González, M.: *Herramienta para la validación de elementos de mejora UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU)*. GRIAL Research Group. University of Salamanca (2017)
6. Anderson, S.P.: *Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences*, Portable Document. Pearson Education (2011)
7. Kveton, P., Jelínek, M., Klimusová, H., Voboril, D.: *Data collection on the Internet: Evaluation of Web-based questionnaires*. *Studia psychologica* 49, 81 (2007)

8. Flavián, C., Gurrea, R., Orús, C.: The effect of product presentation mode on the perceived content and continent quality of web sites. *Online Information Review* 33, 1103-1128 (2009)
9. Huber, W., Vitouch, P.: Usability and accessibility on the internet: Effects of accessible web design on usability. In: *International Conference on Computers for Handicapped Persons*, pp. 482-489. Springer, (2008)
10. Seckler, M., Tuch, A.N., Opwis, K., Bargas-Avila, J.A.: User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field. *Interacting with Computers* 24, 107-118 (2012)
11. Shneiderman, B.: Designing trust into online experiences. *Communications of the ACM* 43, 57-59 (2000)
12. Recabarren, M., Nussbaum, M.: Exploring the feasibility of web form adaptation to users' cultural dimension scores. *User Modeling and User-Adapted Interaction* 20, 87-108 (2010)
13. Pengnate, S.F., Sarathy, R.: An experimental investigation of the influence of website emotional design features on trust in unfamiliar online vendors. *Computers in Human Behavior* 67, 49-60 (2017)
14. Sayago, S., Blat, J.: Some Aspects of Designing Accessible Online Forms for the Young Elderly. In: *WEBIST* (2), pp. 13-17. (2007)
15. Stieger, S., Reips, U.-D.: What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study. *Computers in Human Behavior* 26, 1488-1495 (2010)
16. ISO/DIS: Draft BS ENISO 9241-220 Ergonomics of human-computer interaction. Part 220: Processes for enabling, executing and assessing human-centred design within organizations. (2016)
17. ISO/DIS: Ergonomics of human-system interaction — Part 125: Guidance on visual presentation of Information. (2013)
18. Bevan, N., Carter, J., Harker, S.: ISO 9241-11 revised: What have we learnt about usability since 1998? In: *International Conference on Human-Computer Interaction*, pp. 143-151. Springer, (2015)
19. Bargas-Avila, J., Brenzikofer, O.: Simple but crucial user interfaces in the world wide web: introducing 20 guidelines for usable web form design (2011)
20. Bevan, N., Spinhof, L.: Are guidelines and standards for web usability comprehensive? In: *International Conference on Human-Computer Interaction*, pp. 407-419. Springer, (2007)
21. Seckler, M., Heinz, S., Forde, S., Tuch, A.N., Opwis, K.: Trust and distrust on the web: User experiences and website characteristics. *Computers in Human Behavior* 45, 39-50 (2015)
22. Bargas-Avila, J.A., Brenzikofer, O., Tuch, A.N., Roth, S.P., Opwis, K.: Working towards usable forms on the worldwide web: optimizing multiple selection interface elements. *Advances in Human-Computer Interaction 2011*, 4 (2011)
23. Couper, M.P., Tourangeau, R., Conrad, F.G., Crawford, S.D.: What they see is what we get response options for web surveys. *Social science computer review* 22, 111-127 (2004)
24. Sánchez-Prieto, J.C., Olmos-Migueláñez, S., García-Peñalvo, F.J.: Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers. *Computers in Human Behavior* 55, 519-528 (2016)